Structure, Function and Evolution of Insect Carboxylesterases

A thesis submitted for the degree of Doctor of Philosophy of The Australian National University

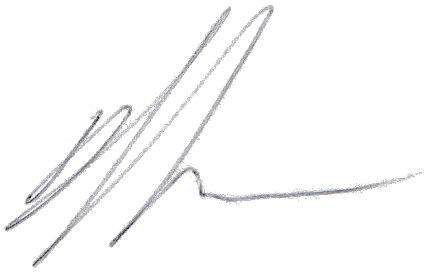
Australian National University

Davis Henry Hopkins

November 2018

Declaration

I declare that the work in this thesis is my own except where it is explicitly stated in the text before each chapter. The majority of the work in this thesis was conducted at the Australian National University under the supervision of Associate Professor Colin Jackson. Some experimental work was conducted at the Commonwealth Scientific and Industrial Research Organization at Black Mountain under the supervision of Dr. John Oakeshott. To the best of my knowledge, the work presented in this thesis has not been submitted as part of any other degrees.



Davis H. Hopkins

November 2018

Acknowledgements

I would like to acknowledge the Australian Government for their support through a Research Training Program Scholarship, and the Australian National University and Research School of Chemistry for their support through the Alan Sargeson Merit Scholarship and the Research School of Chemistry Scholarship.

First and foremost, I would like to thank my supervisor Associate Professor Colin Jackson. It's been a long journey but your open personality, willingness to listen and candid advice have been of immeasurable help. I'll remember fondly the comedically awkward group meetings, as we searched for the right format, and the cunning bartering over board games. I was lucky to have you as a supervisor.

I would also like to thank everyone who I collaborated with, particularly Dr. John Oakeshott, who took me in when I needed a boost, and exposed me to the distinct CSIRO life. Your wisdom, patience and soft-spoken humor were greatly appreciated. To Dr. Rahul Rane, your phylogenetic skills made a complex task simple. To Dr. Chris Coppin and Dr. Faisal Younus, your quick banter, readiness to chat and prompt advice made my time at CSIRO that much nicer.

The Jackson lab has gone through many changes over the years but has always been filled with intriguing, animated people. I am truly grateful for the fun, laughter and absurdity I got to share with many of you. It has been a great place to work and play.

To my family and friends, I owe a huge thank you. From the healthy distractions of frisbee with friends to the warmth and comfort of fun with family, I sincerely appreciate it all.

Finally, I would like to thank my partner, Caroline. For our many adventures, both at home and abroad, for your enthusiasm for life, and for supporting me through all the challenges. You've had to bear a fair bit of my burden, particularly in the final stretch, yet have brought me such joy. I honestly cannot thank you enough.

Abstract

Insect carboxylesterases (CBEs) have proven to be a highly adaptable family of enzymes that has undergone extensive functional diversification and sequence divergence over a short span of evolutionary time. This makes these enzymes ideal examples to explore the evolutionary processes that lead to the unique functions of enzymes. In this thesis I present two such examples.

The first example is addressed in chapters 2 and 3: the evolution of insecticide resistance CBEs. These enzymes are implicated in the most common forms of insecticide resistance, a global issue that threatens both our agricultural productivity and health. While a great deal of work has gone into the identification of insecticide resistance CBEs, there has been little molecular characterization of these enzymes. This is vital to better understand how they function and to allow target-based inhibitor design to combat the resistance they provide. In chapter 2, I describe my attempts to express a large range of insecticide resistance CBEs in *Eschericia coli*. This is a critical first step in the large-scale expression required for crystallization and full characterization. I identified five insecticide resistance CBEs with sufficient expression for crystallization trials. In chapter 3, I describe the crystallization and characterization of one of these CBEs, Cqestβ2[1], which is the most common insecticide resistance CBE in the important disease vector, *Culex quinquefasciatus*. Cqestβ2[1] is the first insecticide sequestration CBE to be structurally characterized. Its structure demonstrates a high similarity to the insecticide target, acetylcholinesterase. Sequence similarity networks of all insect CBEs demonstrated that insecticide resistance CBEs share a level of similarity. This was further emphasized through a structural comparison between Cqestβ2[1] and other insect CBEs. Kinetic characterization of Cqestβ2[1] supported its role in organophosphate resistance via sequestration. Finally, a comparison between Cqestβ2[1] and its naturally occurring isoforms suggests target-based inhibitor design may have broad applicability.

The second example is addressed in chapter 4: the evolution of an odorant degrading enzyme (ODE) from a juvenile hormone esterase (JHE) duplication in *Drosophila melanogaster*. While the evolution of new functions via gene duplication is a widely accepted mechanism, there are relatively few, well characterized examples of this

process. The distinct regulation and substrate specificities of these enzymes also provides a unique opportunity to explore the interaction of both structural and regulatory changes in neofunctionalization. A phylogenetic analysis shows that JHEs have been the template to many distinct functional groups of enzymes. Biochemical comparison reveals sufficient promiscuity in the *D. melanogaster* JHE (DmJHE) to have immediate utility as an ODE. Homology modelling and comparison with known structures of insect JHEs and ODEs revealed similarities and differences that distinguish these groups and suggests key structural changes that explain this example of neofunctionalization.

Finally, in chapter 5, I discuss the significance of my research and the insights that these two examples provide to the process of enzyme evolution. The first, the insecticide resistance CBEs provide a critical example of the early stages of enzyme evolution whereby a promiscuous activity results in a novel function. The comparisons drawn between Cqestβ2[1] and LcαE7, an insecticide resistance CBE from *Lucilia cuprina* that utilizes catalytic detoxification, emphasize distinct strategies through which natural evolution selects for novel functions. The second, DmJHE and DmJHE duplication, provides an example of a later stage in the process of neofunctionalization whereby structural and regulatory changes have resulted in two distinct enzymes with unique functions.

Table of Contents

Abbreviations

AaB1	*Aedes aegypti* B1
Ac-CCE	*Anisopteromalus calandrae* CCE
AChE	Acetylcholinesterase
AgB2	*Anopheles gambiae* B2
BdB1	*Bactrocera dorsalis* B1
CBE	Carboxylesterase
CcαE7	*Ceratitis capitata* αE7
CpCE-1	*Cydia pomonella* CE-1
Cqestα2[1]	*Culex quinquefasciatus* estα2[1]
Cqestβ2[1]	*Culex quinquefasciatus* estβ2[1]
DmEST6	*Drosophila melanogaster* Esterase 6
DmJHE	*Drosophila melanogaster* juvenile hormone esterase
DmJHEdup	*Drosophila melanogaster* juvenile hormone esterase duplication
EC	Enzyme commission
Est-A	Esterases-A
Est-B	Esterases-B
Est-C	Esterases-C
gor	Glutaredoxin reductase
Hax001D	*Helicoverpa armigera* 001D
IUBMB	International Union of Biochemistry and Molecular Biology
JH	Juvenile hormone
JHE	Juvenile hormone esterase
MpE4	*Myzus persicae* E4
MpFE4	*Myzus persicae* FE4
Nl-EST1	*Nilaparvata lugens* EST1
ODE	Odorant degrading enzyme
OP	Organophosphate
RmEST9	*Rhipicephalus microplus* EST9
RMSD	Root mean square deviation
Seq ID	Sequence identity
SP	Synthetic pyrethroid
SSN	Sequence similarity network
S-tag	Solubility tag
T7 RNAP	T7 RNA polymerase
trxB	Thioredoxin reductase

Chapter 1. General Introduction

1.1. Overview

This section will introduce carboxylesterases (CBEs) highlighting their general catalytic mechanism and the important roles they play in biology. I will also discuss their current classification with a particular focus on insect CBEs. Commonalities in their structural and catalytic features will be presented before briefly introducing the process of enzyme evolution and how it relates to insect CBEs.

1.2. CBE's diverse roles and functions

CBEs are enzymes that catalyze the hydrolysis of a carboxyl ester through the addition of water, thereby converting the ester into an alcohol and a carboxylic acid (**Figure 1.1**) (1). They are widely distributed in all forms of life and are critical in a large range of biological processes including hormone regulation, neurotransmission, digestion and xenobiotic metabolism (2–5). Much of the research on these enzymes has focused on their ability to degrade xenobiotics, whether this be in humans, relating to drug activation and degradation, or in insects and bacteria, relating to the evolution of pesticide resistance (6–10). The latter presents an interesting case where enzymes have evolved over a relatively short period of time providing a useful example of enzyme evolution that gives insight into the process of natural evolution and how we can better engineer enzymes (11–14).

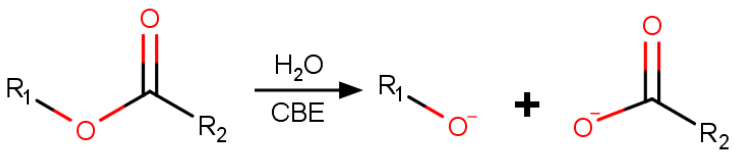


Figure 1.1. CBEs catalyze the hydrolysis of carboxyl esters through the addition of water converting them to free alcohol and carboxylate molecules.

1.3. Classification of CBEs

Correct classification of enzymes is critical in better understanding how they relate to each other, interact and in allowing us to infer the properties of newly discovered enzymes without experimentation, which is difficult for all of the many thousands of

sequences obtained through genomic and metagenomic sequencing. There have been many attempts to create an ideal classification system for CBEs but as our interest and understanding of these enzymes has increased faults in each have emerged (15). This has resulted in many different systems being simultaneously used based on both convenience and relevance to individual studies and fields. The advantages and disadvantages of each will be discussed below.

Inhibition classification- In 1953, Aldridge proposed one of the first systems of classification that relied solely on the esterase's interaction with insecticidal organophosphate esters (OPs) (16, 17). This classification divided esterases into three groups: esterases A (Est-A), capable of hydrolyzing OPs; esterases B (Est-B), inhibited by OPs; and esterases C (Est-C), which do not interact with OPs (16). These groups were seen to be too broad due to a majority of esterases falling into the Est-B category even though they had differing functions (18). An expansion of this system was later proposed that used a range of inhibitors: sulfhydryl reagents (mainly p-chloromercuribenzoate), OPs (paraoxon, amongst others) and carbamates (exclusively eserine) (19). This separated the esterases into four distinct classes:

1. *Arylesterases,* which preferentially hydrolyze aromatic esters and are only inhibited by sulfhydryl reagents.
2. *CBEs*, which preferentially hydrolyze aliphatic esters and are exclusively inhibited by OPs.
3. *Cholinesterases*, which react with choline esters at a higher rate than both aliphatic and aromatic esters and are inhibited by both OPs and carbamates.
4. *Acetylesterases,* which preferentially hydrolyze aromatic esters but are not inhibited by any of the tested inhibitors (19).

Classification based on inhibition was able to distinguish esterases into a small number of groups relevant to the enzymatic interaction with insecticides but as our interest in this group of enzymes progressed more detailed forms of classification were necessary to better distinguish and understand the diverse roles of this complex family.

Enzyme Commission classification- In 1956, the International Union of Biochemistry (now the International Union of Biochemistry and Molecular Biology (IUBMB)) established the International Commission on Enzymes to address the difficulties in

naming enzymes (20). The number of enzymes being discovered was rapidly increasing and due to the lack of a naming system inconsistencies were a major issue (21). In 1961, the first report was presented establishing the foundations for the enzyme commissions nomenclature of enzymes that are still used today (21). This was based on the principle that enzymes should have names indicating the reactions they catalyze. The role of updating and maintaining the enzyme nomenclature has been taken over by various committees resulting in a progressively more comprehensive set of recommendations and supplements for the suggested enzyme nomenclature (22).

The Enzyme Committee classification system classifies enzymes using a series of four code numbers separated by points and preceded by the letters EC (Enzyme Commission):

i) The first number divides enzymes into 6 classes based on their type of reaction (*e.g.* EC3 refers to enzymes that hydrolyze bonds using water)
ii) The second number defines the nature of the chemical bond acted upon (*e.g.* EC3.1 indicates reaction with ester bonds)
iii) The third number relates to the broader nature of the substrate (*e.g.* EC3.1.1 indicates reaction with carboxylic ester bonds)
iv) The fourth number refers to a specific class of substrate (*e.g.* EC3.1.1.8 indicates reaction with an acylcholine ester)

As shown above, enzymes that hydrolyze carboxylic esters can be defined by the first three numbers and are classed as EC3.1.1 This group is then divided into 103 subgroups related to the specific class of substrate that the enzymes act upon (22). This system offers far greater distinction than the inhibition classification system and is sufficient for many other enzyme families (23). However, due to many CBEs possessing broad and overlapping substrate ranges they may appear in the same EC group while being from distinct species and functioning in very distinct roles physiologically. For example, odorant degrading enzymes (ODEs), like Esterase 6 from *Drosophila melanogaster*, and metabolic enzymes, like LcαE7 from *Lucilia cuprina*, both react with fatty acid methyl esters and would be classed as EC 3.1.1.1. However, the former is involved in the olfactory system and the latter in digestion (5, 24). Another critical issue in this classification emerges from the inherent structural plasticity and tendency for gene duplication in the CBE gene family (18, 25, 26). This,

combined with the rapid generation time of insects, allows CBEs to quickly adapt and evolve activities with different substrates (11, 27).

Time of discovery, substrate and electrophoretic mobility criterion– Many species-specific nomenclatures have emerged based on easily identifiable, enzymatic features including a combination of time of discovery, substrate preference and protein gel mobility. Some aphid CBEs (E1-E7) are named solely based on their mobility after native electrophoresis (28). *Drosophila* CBEs and isozymes have generally been classified by both their electrophoretic mobility and their preferential hydrolysis of the artificial substrates α- and β-naphthyl acetate (α-esterases and β-esterases, respectively) (29, 30). *Culex* CBEs use the criteria of *Drosophila* and extend it. First the substrate preference and electrophoretic mobility is indicated as in *Drosophila* (*e.g.* Estβ2) but then, once the CBEs are characterized at the nucleotide level, a superscript number is added allowing a distinction between CBEs that differ at the nucleotide level but not electrophoretically (*e.g.* Estβ2^{1} and Estβ2^{2}) (31, 32). The convenience of this nomenclature is clear as in theory a CBE can be identified through a few simple tests. Unfortunately, each test has a number of problems in its use as an identifier. For example, a significant difference in nucleotide sequence is possible without affecting electrophoretic mobility and is thus hidden by using this technique (15). Also due to the generally broad substrate range of CBEs, the preferential activity against only two substrates is a poor indicator of their native substrates and can be misleading (*e.g.* an α-esterase and β-esterase may act on similar native substrates) (32, 33).

Phylogenetic criterion- As more CBEs have been characterized, and genomic data for a large number of species has become available, attempts at classifying CBEs through the use of phylogenetic analysis has become more common (30, 34). There have been several attempts to classify insect CBEs phylogenetically and, as each new insect genome is identified, refinements to the phylogeny are suggested. One of the first comprehensive classifications was suggested by Oakeshott *et al.* (2005) and used sequences from the genomes of *Drosophila melanogaster* and *Anopheles gambiae* and various characterized CBEs from other insects (30). This divided CBEs into fourteen clades (A-N) based on largely monophyletic groups and named based on function. These clades were grouped into three broad classes based on function: dietary/detoxification, hormone/semiochemical processing and neuro/developmental

functions. As more phylogenies have been proposed the number and nature of the clades has changed, yet the classes have persisted.

In 2010, the most comprehensive phylogeny produced so far was made using the genomes of seven insect species from the orders Diptera, Hymenoptera and Coleoptera, which resulted in a slight re-working of the definitions for the clades, particularly those in the dietary/detoxification class (35). This was further expanded through the addition of CBEs from the genomes of *Bombus terrestris* and *Bombus impatiens* in 2015, yet the clades were unchanged (**Figure 1.2**) (36). A description of this phylogenetic classification follows.

The dietary/detoxification class contains three clades: clade A, hymenopteran xenobiotic metabolizing enzymes; clade B, α-esterase type enzymes, generally microsomal; and clade C, unknown functions. These CBEs generally have broad activities making their classification difficult. A number of other phylogenies, utilizing Lepidoptera and Hemiptera, suggest greater complexity exists in this class and that more clades may be required to fully describe the diversity of CBEs (37–40).

The hormone/semiochemical processing class contains four clades: clade G, lepidopteran-type JHEs; clade D, integument esterases; clade E, secreted β-esterases; and clade F, dipteran-type JHEs. The functional roles of these clades are better defined and are thus more consistently supported with the exception of clade D and clade E. Previous phylogenies have suggested that these clades may contain multiple monophyletic groups (39, 41). As more insect CBEs are characterized further distinction within these broader clades may be determined.

The neuro/developmental class is the most consistently supported and contains six clades: clade H, glutactin enzymes; clade J, acetylcholinesterases (AChEs); clade N, neurotactins; clade I, unknown function; clade M, gliotactins; and clade L, neuroligins. The proteins in these clades generally have critical and distinct functions resulting in the higher support of these groupings. While they all have high sequence homology to insect CBEs some clades contain non-catalytic proteins (35).

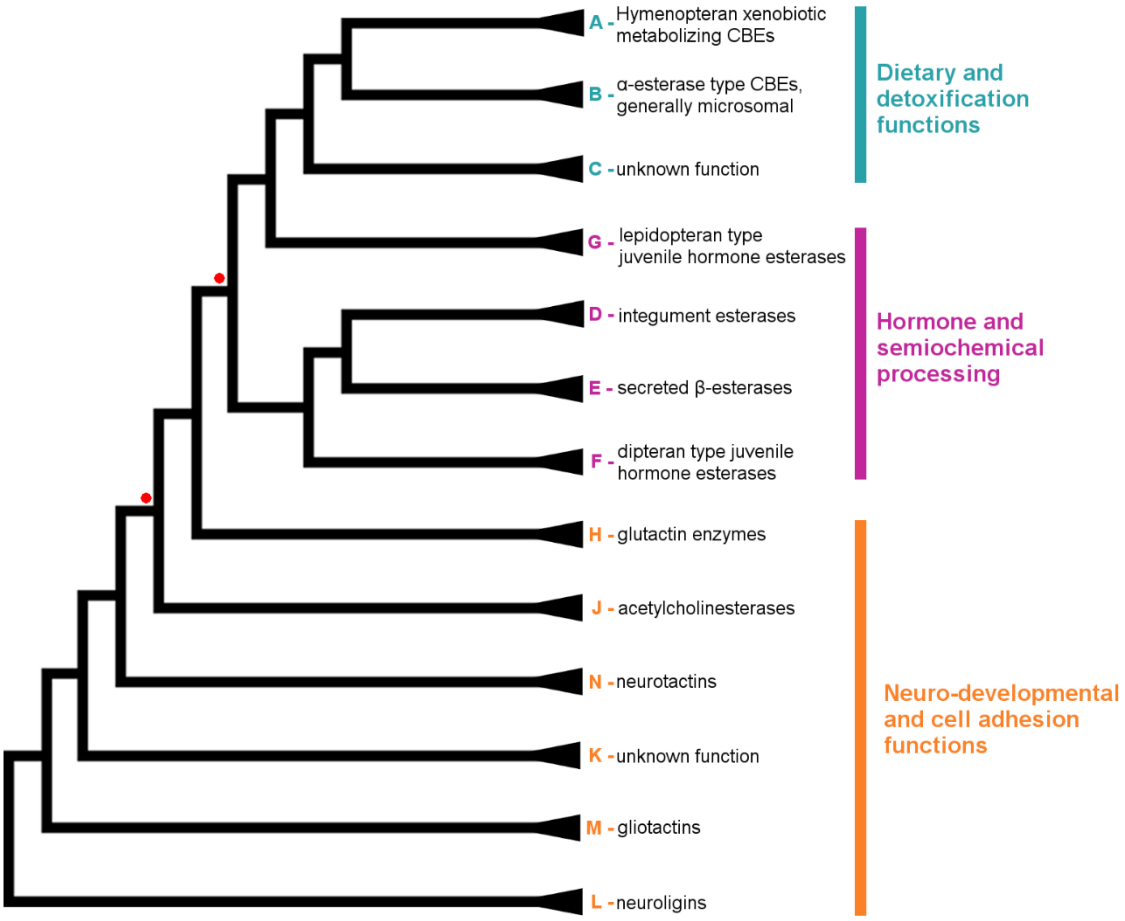


Figure 1.2. A diagrammatic representation of the phylogenetic clades proposed by Oakeshott et al. (2010) colored based on their suggested functional classes. Branches with percentage bootstrap confidence values above 50% are marked at nodes by a red dot (35).

The phylogenetic criterion for insect CBE classification has proven to be the most detailed in describing the complexity of both evolutionary relationships and functional groupings present in the CBE family. One of the key issues with this criterion is the lack of CBE characterization, particularly in the CBEs in the dietary/detoxification class and clade E of the hormone/semiochemical processing class. Another key issue is the occurrence of stochastic error. The generation of a phylogenetic tree requires careful choice of CBEs to be included in the analysis to best generate monophyletic groups with high confidence. Unfortunately, even in the most comprehensive trees this has resulted in some CBEs being excluded that in other trees define unique monophyletic groups. To date, there has not been a full representation of the diversity and groupings present in the CBE family.

1.4. α/β hydrolase fold

CBEs belong to the carboxyl/cholinesterase gene family (Pfam PF00135) within the α/β hydrolase fold superfamily (42). The α/β hydrolase fold was determined in 1992 through the significant similarities found between the structures of five hydrolytic enzymes with varying functions: dienelactone hydrolase, haloalkane dehalogenase, wheat serine carboxypeptidase II, AChE and a lipase (43). In 1995, the ESTHER database was created to gather biochemical, pharmacological and structural data for annotated genes and protein sequences with an α/β hydrolase fold (44). To date, at least 489 different proteins have been structurally determined with at least one structure present in 114 distinct, functional families (44). This superfamily contains proteins with a wide variety of catalytic functions including proteases, lipases, CBEs, peroxidases, dehalogenases, as well as a number of non-catalytic functions including gliotactins, neurotactins, neuroligins and more (44, 45). While members of the superfamily adopt the same fold they can share very little sequence identity demonstrating the plasticity in this fold and exemplifying its evolvability (46).

The canonical α/β hydrolase fold provides a stable scaffold for a wide variety of functions and is composed of a mostly parallel, eight stranded β sheet, surrounded on each side by a total of six α helices (**Figure 1.3**) (45). For the catalytic members of this superfamily the catalytic residues are closely associated with this core fold and are thus highly conserved (45). These most often consist of a nucleophile (serine, cysteine or aspartic acid) positioned at a sharp turn after β-strand 5, an acidic residue positioned after β-strand 7 and an absolutely conserved histidine residue positioned after β-strand 8, all of which form a catalytic triad (**Figure 1.3**) (45). The sharp turn that the nucleophile is present on is often referred to as the ‘nucleophile elbow’ and adopts an energetically unfavorable main chain torsion angle (45). The geometry of the nucleophile elbow also imposes steric restrictions on the surrounding residues and contributes to the formation of the oxyanion-binding site (the oxyanion hole), a critical feature of the active site. This oxyanion hole is formed by two or three backbone nitrogen atoms that stabilize the developing negatively charged transition state that occurs during hydrolysis (45). The binding of the substrate molecule is defined by a range of loops, helices and strands that form subdomains that extend above the

catalytic residues and widely vary between enzymes in the superfamily (45). For CBEs, the substrate binding site is often defined by two subdomains that stabilize the esters acyl and alcohol groups into distinct pockets, positioning the carbonyl group for ideal interaction with the catalytic residues and stabilization of the transition state by the oxyanion hole (45).

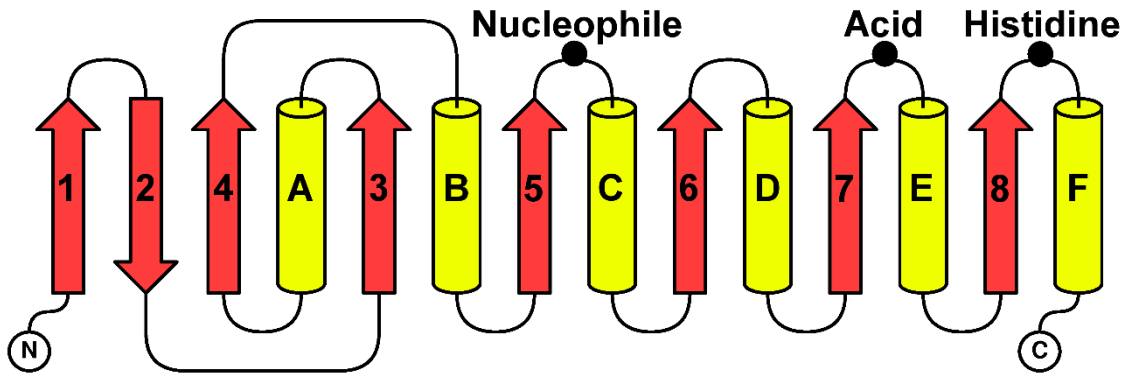


Figure 1.3. The secondary structural topology map of the canonical α/β hydrolase fold (45). β-strands are numbered and shown in red and α-helices are lettered and shown in yellow. The catalytic residues are represented by dots and labelled.

1.5. The catalytic mechanism of CBEs

The catalytic mechanism of CBEs is shared amongst other members of the α/β hydrolase fold family and is through base-mediated hydrolysis, which relies on a catalytic triad (47–52). Interestingly, catalytic triads are common among a number of other protein fold superfamilies and are thought to be examples of convergent evolution (53). Using CBEs as an example, the hydrolysis proceeds through a number of steps (**Figure 1.4**). First the catalytic nucleophile (commonly serine) is activated through interaction with a histidine residue, which itself is positioned to be activated by an acidic residue (**Figure 1.4A**) (47–52). This nucleophile covalently binds with the substrates carbonyl group, releasing the double bond and resulting in the formation of a negatively charged, tetrahedral transition state (**Figure 1.4B**) (47–52). The transition state then spontaneously collapses, releasing the ester's alcohol group and forming an acyl-enzyme link (**Figure 1.4C**) (47–52). Finally, a water molecule interacts with the acyl-enzyme complex reforming a tetrahedral intermediate before spontaneous collapse and release of the acidic product leaving the nucleophile free for further

reaction (**Figure 1.4D-F**) (47–52). The water molecule is thought to either be activated by the histidine and acid residues of the catalytic triad or other acidic residues in the active site (47–52).

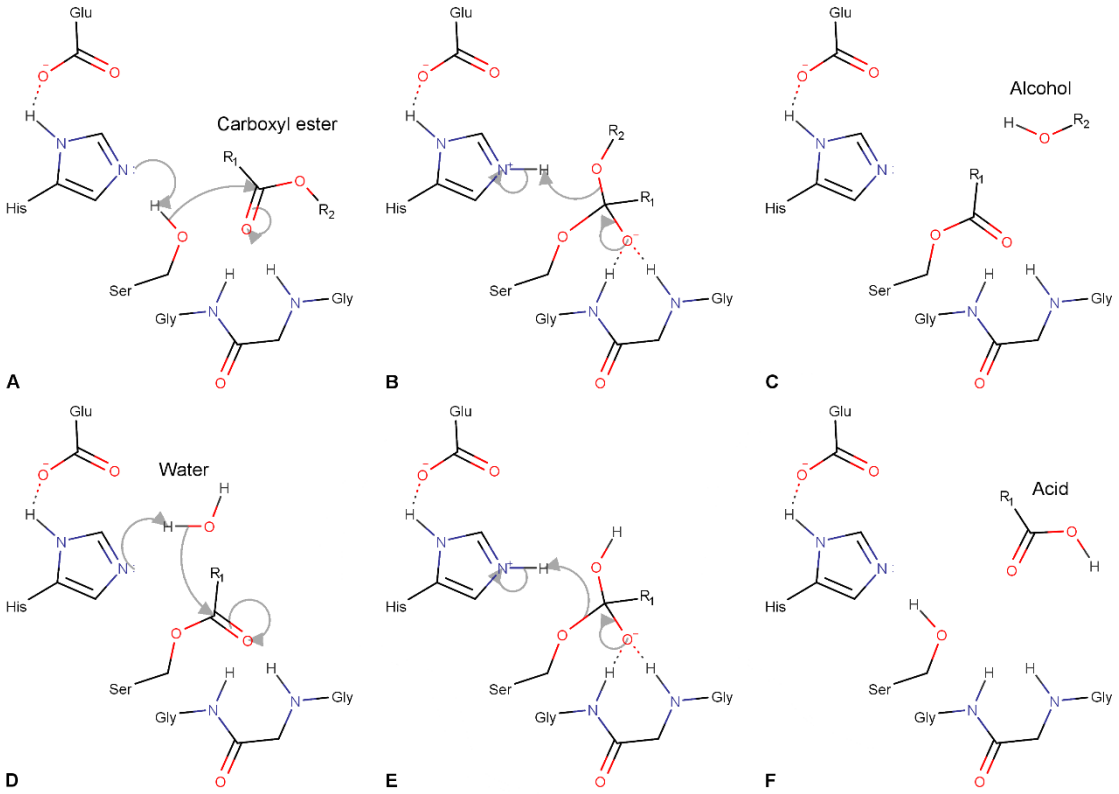


Figure 1.4. The mechanism of a CBE reacting with a carboxyl ester. (**A**) The catalytic serine is activated by the coordinated histidine to react with the carboxyl ester resulting in a tetrahedral intermediate (**B**), which is stabilized by the glycine residues of the oxyanion hole. This collapses releasing the alcohol group, regenerating the histidine and resulting in an acyl-enzyme intermediate (**C**). (**D**) A water molecule is activated by the catalytic histidine to react with the acyl-enzyme intermediate resulting in a new tetrahedral intermediate (**E**). This collapses releasing the acid group from the catalytic serine and regenerating the catalytic triad (**F**).

1.6. Enzyme evolution

Historically, it was thought that one of the key features that made enzymes such good catalysts was their high specificity towards one substrate (54). However, as our knowledge of enzymes has increased, it has become clear that many, if not most, enzymes are capable of reacting with substrates that they were not evolved to specifically react with (54). In 1976, Jensen was one of the first to suggest that this

promiscuity is key to the evolution of enzymes (55). He argued that the earliest forms of life had a small number of genes encoding enzymes with broad specificities allowing function with a wider range of substrates. Duplication of these genes would have created redundancies in their encoded enzyme's activity allowing the accumulation of mutations to some of the duplicated enzymes that specialized them towards specific reactions. The increased catalytic efficiency of these enzymes would improve organismal fitness and thus be selected for. This process of duplication, mutation and selection has formed the basis for our understanding of enzyme evolution (56).

There are a number of different ways to describe the ability of certain enzymes to react with multiple substrates, all of which have implications for the evolution of new enzymes. Broad-specificity refers to enzymes that have evolved to transform a range of substrates with similar efficiency. This differs to substrate ambiguity, which refers to enzymes capable of reacting with substrates they would not normally encounter but that possess similar structures to their native substrate. Enzyme promiscuity refers to an enzyme displaying activity that it did not evolve for and that is not a part of the organism's normal physiology (54, 57–59).

The ability of certain enzymes to catalyze multiple reactions must be inherent in their structure. This can be due to a number of features in the active site including the existence of different subsites, which could result in a non-specific binding pocket that provides a broad-specificity but low catalytic efficiency with any one substrate (54). Multiple reactions can also be made possible through conformational diversity, whereby the inherent plasticity in an enzyme's structure enables it to adopt alternate conformations that favor promiscuous activity (56, 60, 61). This is highlighted by Campbell *et al.* in their study of the laboratory evolution of a *Pseudomonas diminuta* phosphotriesterase to an arylesterase (62). In this, they demonstrated that bi-functional intermediates were surpassed through mutations that favored conformations related to the arylesterase activity. Interestingly, one such bi-functional intermediate has recently been observed in the naturally occurring αE7, an insecticide resistance CBE from *L. cuprina* (63). The promiscuous activity against OPs is restricted to a conformation that is sampled infrequently, restricting its catalytic potential.

As there is no ‘fossil record’ of ancestral proteins, our understanding of the mechanisms and mutational pathways that led to today’s enzymes rely on comparisons of phylogenetically related enzymes, experiments that computationally can infer ancestral proteins or through laboratory directed evolution, studies that mimic the process of evolution over shorter periods of time (64). Phylogenetic studies are complicated by the fact that enzymes with distinct functions that are closely related, phylogenetically, may differ from 30% to 80% in sequence with an unknown proportion of these changes relating to “neutral drift” rather than change of function mutations (54). Thus, directed evolution experiments have formed the major basis for our understanding of the mechanisms behind enzyme evolution (56).

For an enzyme to develop a new function there are a number of requirements. Firstly, the evolving activity must give an immediate physiological advantage to allow for selection (54). Secondly, once this novel function is physiologically relevant, further improvement must be possible through limited mutation due to the restrictive mutational rate of living organisms (54). Directed evolution experiments have suggested that the mutational pathway from a native function to a novel one may proceed within the bounds of two extreme pathways (**Figure 1.5**) (65–68). The first extreme is thought to be the main route and is through weak negative trade-off, or the convex route, where mutations greatly increase the novel activity while slightly reducing the native activity (54). The earliest mutations towards novel activity often occur outside the protein scaffold or active site (60). While the mechanism for novel activity improvement from this is still unclear, it is likely to be related to a shift in conformation that better accommodates the novel activity. The location of these mutations is also least likely to affect the native activity and is a key feature of the enzyme’s perceived plasticity. This route is likely to result in a generalist enzyme before a sudden dramatic reduction in native activity resulting in specialization towards the novel activity. The other extreme is through strong negative trade-off, or the concave route, where a slight increase in novel activity comes at a great cost to the native activity. Unlike the other extreme this one would require a gene duplication so that the physiologically relevant native activity can be retained as evolution occurs (67, 69–73).

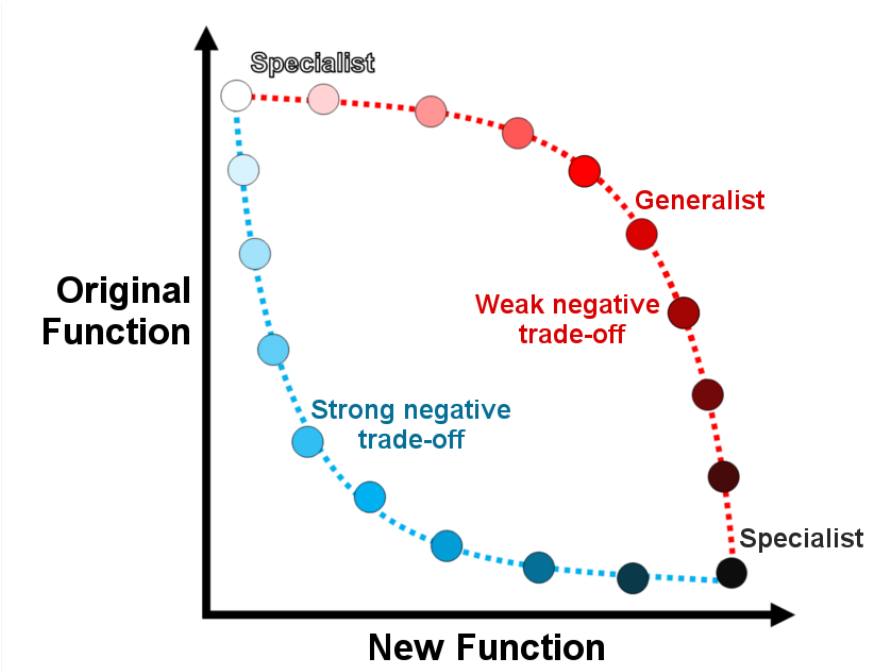


Figure 1.5. The two extremes of the mutational pathways possible for an enzyme to develop a new function. The weak negative trade-off pathway (the convex route) involves first a gradual shift from the specialist enzyme in the original function (white) with large increases in the new function. It then transitions through a generalist enzyme (red) that is equally capable of both the original and new function before a dramatic decrease in the original function to result in a specialist in the novel one (black). The strong negative trade-off pathway (the concave route) transitions through an enzyme weak in both and likely requires a gene duplication to maintain the original function while a duplicate enzyme evolves the new one. Adapted from references (54, 68).

There are a number of other important factors that are not addressed by the above description. First is the effect that mutations have on the stability of enzymes. Mutations, in particular those related to novel functions, have most often been shown to be destabilizing (74–76). Thus, when evolving towards a new activity, enzymes often become destabilized resulting in reduced expression through misfolding and aggregation. This is a major constraint of protein evolution (77). Destabilizing mutations must also be compensated for through the addition of stabilizing mutations that do not substantially affect the novel activity. Chaperones can also provide a buffering effect, allowing enzymes with lowered stability but improved activity to still be selected for (78). Another important factor to consider is that of epistasis, where the effects of a particular mutation may change depending on those that came before it (79). Considering only two mutations there are four possible outcomes (79):

1. No epistasis - where the addition of each mutation alone has the same magnitude of effect upon their combination.
2. Magnitude epistasis – where the combination of each mutation results in an increase of effect greater than each alone.
3. Sign epistasis – where one mutation is deleterious alone but when combined with an advantageous mutation also provides an improvement
4. Reciprocal sign epistasis – where both mutations are deleterious alone but when combined have an advantageous effect.

This effect is important to consider when tracing the evolution of novel enzyme functions (14, 77)**.**

1.7. Evolution of the CBE multigene family

Due to the diversity of biological esters and their importance in life there appears to have been a great deal of diversification of esterases in a number of protein fold superfamilies (12, 15, 80, 81). Early evolution in prokaryotes required the capability to hydrolyze thio-, phospho- and carboxyl esters (43). As organisms grew more complex and required more processing of biological esters, esterases diversified accordingly. Eukaryotes developed a range of novel hormonal, neuronal and metabolic characteristics many of which utilized esters and thus required further control over the metabolism of these molecules (82, 83). Interestingly, while esterases in other superfamilies played a role in this, the enzymes from the α/β hydrolase fold superfamily proliferated and diversified far more, resulting in a number of new multigene families associated with distinct roles (84, 85). The CBE gene family was one of these, which has continued to diversify to accommodate the increasing biological complexity (86).

CBEs are found in animals, plants and (to a limited degree) in bacteria, indicating that they likely evolved in ancient bacteria but proliferated predominantly in eukaryotes (87). There are relatively few members of the CBE gene family in lower eukaryotes such as yeast, suggesting proliferation was predominantly in the higher eukaryotes (88–90). Oakeshott *et al*. developed the most comprehensive phylogeny of the CBE multigene family (86). This demonstrated that a number of lineages differentiated

before the prokaryote/eukaryote split and that three of the key splits in the phylogeny coincide with the three major divisions of the prokaryota (86). In this phylogeny there are a number of shared groups across the different classes of the Metazoa, in general, these include specialist proteins such as cholinesterases. Other major radiations clearly occur after the separation of the major classes of the Metazoa and have resulted in class specific groups (86).

One of the key features revealed by their phylogeny was the rapid evolution of paralogues in insect CBEs with paralogues within sub-lineages that separated in the last 50 million years having as little as 60% sequence identity (86). This demonstrates the ability for this gene family to tolerate rapid sequence changes, enabling them to adopt novel functions and explaining the ongoing diversification of this family (86). The phylogeny also demonstrated that CBEs with the highest identity were generally physically co-located on chromosomes (86). This confirms that local amplification through gene duplication has played a critical role in the evolution of this gene family and explains the increased copy number of CBEs in insects (25, 26). Specific examples of this are the α-cluster of CBEs within the higher Diptera in which a varying number of CBEs are present in different insect organisms but all diversification can be traced through robust phylogeny to a series of single gene duplication events (91, 92). These duplication events would have allowed individual genes to accumulate mutations through both weak and strong negative trade-offs and may explain the fast rates of change observed. Interestingly, CBE genes within the α-cluster of *Drosophila* were found to be lost quickly during cluster evolution due to nonsynonymous mutations that occur with a frequency equal to or greater than synonymous mutations (92, 93). This demonstrates that qualitative shifts in function can occur through widespread changes in protein sequence over short periods of evolutionary time.

The widespread use of insecticides and selection for insecticide resistance has provided a unique example of the early stages of enzyme evolution in insect CBEs. Two of the major types of insecticides, carbamates and OPs, target the CBE, AChE. Due to the high rate of gene duplication present in the α-cluster of CBEs, their role in metabolism and their shared mechanism with AChE, this group of CBEs were ideally placed for the evolution of insecticide resistance (11, 94–96). The majority of insecticide resistance CBEs known are from this group (95).

1.8. Focus of thesis

Our interest in insecticide resistance CBEs has resulted in many studies focusing on their expression within insects and in identifying potential insecticide degrading mutations. Surprisingly, there are relatively few examples of these CBEs being biochemically characterized and even fewer that have been structurally characterized. Such studies are key in better understanding the interaction of CBEs with insecticides. In the second chapter of this thesis I will further introduce insecticide resistance CBEs and detail my attempts to express a number of insecticide resistance CBEs, from a diverse range of insect orders, in *Escherichia coli*. This is a critical first step in both biochemical and structural characterization and has enabled two new insecticide resistance CBE structures to be solved. The third chapter contains a first authored journal article focusing on one of these insecticide resistance CBEs, Cqestβ2[1], from the disease vector, *Culex quinquefasciatus*. This CBE is both structurally and biochemically characterized providing a molecular-level understanding of the sequestration mechanism shared by the majority of insecticide resistance CBEs. The relationship between all insecticide resistance CBEs is investigated through a sequence similarity network of all insect CBEs. This reveals that the insect CBEs are a more diverse family than previously thought, with a number of new subfamilies suggested and functionally annotated.

The majority of biochemical and structural studies on enzyme evolution rely on laboratory-simulated evolution. However, to better understand the evolution of enzymes it is important to also study naturally occurring examples that have evolved over longer periods of time. The fourth chapter of this thesis contains a first authored journal article that focuses on one such example in *Drosophila melanogaster*; the evolution of JHE duplication (DmJHEdup), an ODE, from DmJHE, an enzyme critical in hormone metabolism. While DmJHEdup is the result of a duplication of DmJHE, its function, localization and expression have been shown to differ dramatically. The similarities and differences of these two enzymes is explored through biochemical analysis with a wide range of esters and structural modelling. This, along with a phylogenetic analysis including other insect JHEs and ODEs, reveals a potential

justification for this evolution that provides insights into the natural evolution of novel enzyme functions.

Chapter 2. Expression of Insecticide Resistance Carboxylesterases

2.1. Introduction

2.1.1. Insecticides

Insecticides are critical in both maintaining sufficient agricultural productivity and fighting insect-borne diseases (15, 97–99). However, the prevalence of insecticide use over the past 60 years has produced an excessive pressure on insects to evolve resistance (11, 95, 100, 101). Thus, the number of insect species identified with resistance has been rapidly increasing, reaching greater than 580 species in 2015 (102). To combat this, a large amount of research has gone into the development of novel insecticides (103–106). So far, insecticides utilizing greater than 25 modes of action and including at least 55 different chemical classes have been produced (102). The five major chemical classes are organochlorines, carbamates, synthetic pyrethroids (SPs), neonicotinoids, and OPs (**Figure 2.1**) (103, 107, 108). These major classes all target the insect nervous system, disrupting its normal function by interacting with one of the following targets: voltage-gated sodium channels; AChE; and nicotinic acetylcholine receptors (103, 107, 108).

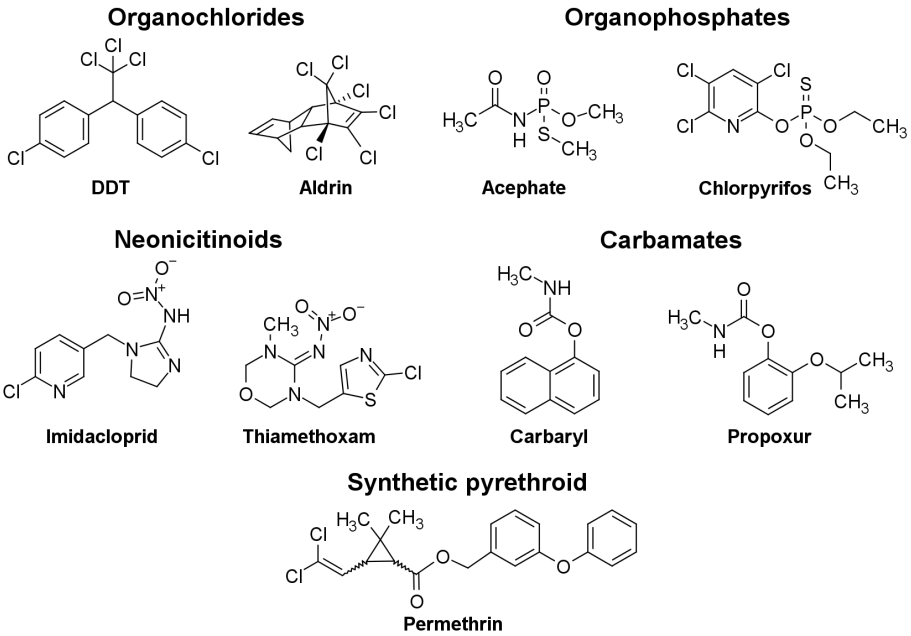


Figure 2.1. A selection of the most commonly used insecticides from the five major classes of insecticides (103, 107, 108).

2.1.2. OP insecticides

OP insecticides are among the most widely used classes of insecticide (102, 109, 110). They (along with carbamates) target AChE, which is expressed in neuromuscular junctions and chemical synapses and hydrolyzes the neurotransmitter, acetylcholine, terminating signal transduction (111–114). This process is essential for the normal physiology of insects and is thus critical for their survival (115–117). The reaction between AChE and OPs results in a covalent linkage between the OP and AChE, which irreversibly inhibits AChE resulting in insect paralysis and eventual death (115–117).

OPs can be described as a group of phosphoric acid ester compounds with three ester bonds (**Figure 2.2**). The first ester bond generally contains a leaving group which consists of an electron withdrawing group that promotes its release during the first step of the reaction between the OP and the catalytic serine of AChE (**Figure 2.3**). The other two ester bonds generally consist of short alkyl side chains (*o*-methyl or *o*-ethyl) (**Figure 2.2**). The OP double bonded oxygen is often replaced by a sulfur atom in commercial synthesis, resulting in a thion form of the OP (**Figure 2.2**) (118). During uptake in insects, native P450 monooxygenases convert this sulfur into an oxygen resulting in the oxon form of the OPs, which are generally stronger inhibitors of AChE (118–120).

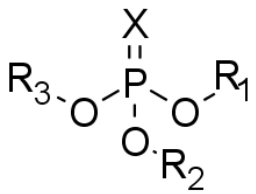


Figure 2.2. The template for an OP structure where X is either an oxygen or sulfur atom, R_1 and R_2 are either *o*-methyl or *o*-ethyl alkyl side chains and R_3 is an electron withdrawing leaving group.

The reaction between AChEs and OPs proceeds via a catalytic triad and shares the first steps with the base-mediated CBE hydrolysis mechanism (**Figure 2.3**) (30, 114). First, the catalytic serine undergoes a nucleophilic substitution at the OP double bonded oxygen releasing a leaving group ester (**Figure 2.3**). Unlike the reaction between a CBE and an ester the covalent enzyme-OP complex formed is very stable

(121–125). The reaction can then proceed via two pathways: a water molecule may hydrolyze the complex reforming the catalytic serine; or through a process called aging, a water molecule may undergo nucleophilic substitution with one of the phosphoryl esters of the complex resulting in an irreversibly inhibited complex (**Figure 2.3**) (30, 114). Due to the stability and steric constraints imposed by the tetrahedral serine-OP complex each pathway proceeds very slowly rendering the enzyme catalytically inactive (121–125).

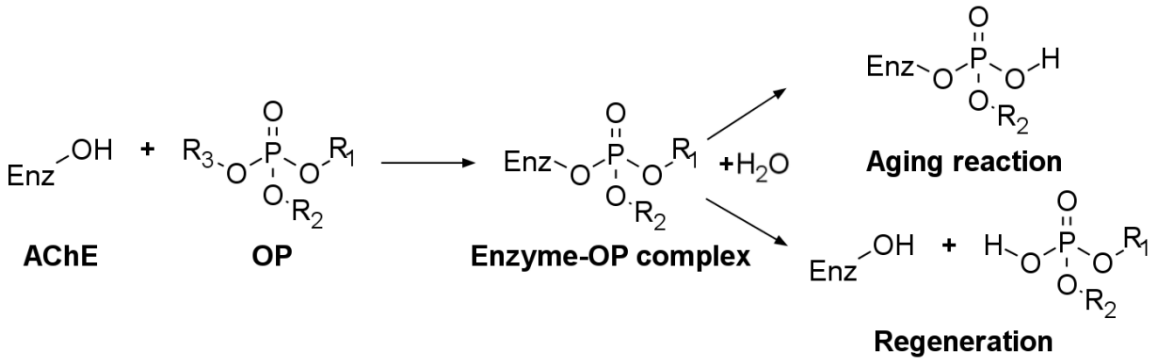


Figure 2.3. The basic reaction mechanism between AChE and an OP showing the two possible pathways from an Enzyme-OP complex.

2.1.3. Insecticide resistance mechanisms

Interestingly, the wide variety and large number of insecticides are countered by relatively few insecticide resistance mechanisms (25, 126–128). This highlights the importance of better understanding these mechanisms to either avoid or directly respond to their complicating effect and thus improve insecticide efficacy. The two main types of insecticide resistance mechanisms are target-site modifications and enzyme-based, metabolic resistance (126, 129–131). Target-site modifications cause a mutation in the target of the insecticide that prevents its intended interaction (132, 133). In the majority of insect targets, insensitivity is caused through very few mutations that are consistent amongst diverse insect species (132–136). There are only three major targets for the majority of conventional insecticides: γ-aminobutyric acid receptor subunit (RDL); a voltage-gated sodium channel (PARA); and AChE (132–137). In these, only one mutation has been found in RDL, two in PARA and three in AChE (132–137). While this form of resistance has been found in many species, the more common mechanism is through metabolic resistance (35, 137, 138). This mechanism relies on the promiscuous activity of detoxification enzymes that degrade

the insecticide before they can reach their target (35, 137, 138). There are three major classes of detoxification enzyme utilized in insecticide resistance: mixed-function oxidases, glutathione S-transferases, and CBEs (94, 126, 139–143).While all have been shown to play significant roles in different organisms, CBEs are the most widespread class of detoxification enzyme and are the most closely associated with the most common forms of insecticides: OPs, carbamates, and to a lesser extent, SPs (144–148).

2.1.4. CBE-mediated metabolic resistance

CBE-mediated metabolic resistance has been identified in many insect species across a large number of insect orders (15, 40, 149, 150). The most common form is through insecticide sequestration, which involves weak promiscuous activity allowing a slow hydrolysis of the insecticide that results in an essentially stoichiometric, covalent sequestration of the insecticide (10, 127, 151, 152). For this mechanism to be effective it requires overexpression of the insecticide resistance CBE to accommodate the insecticide. This mechanism is sometimes referred to as the quantitative resistance mechanism due to its reliance on this overexpression, which is commonly achieved through tandem amplifications of the insecticide resistance CBE genes (96, 153, 154). Aphids and culicine mosquitos are the most well studied organisms that utilize this mechanism, where tandem amplification has been shown to generate up to 200 or more copies of the resistance gene (25, 155–158).

The less common form of CBE-mediated resistance is through catalytic detoxification (10, 159, 160). This involves the accumulation of mutations that increase the hydrolysis of insecticides (160, 161). There are significantly fewer examples of CBEs that utilize this mechanism (5, 63, 96). The most well studied is the α-esterase from *L. cuprina,* LcαE7 (5, 63, 161–163). In this enzyme, a glycine residue from the oxyanion hole of the active site is mutated to an aspartate residue (G137D), which is thought to act as a general base that activates a water molecule enabling complete hydrolysis of OPs (63, 163, 164). This has been shown to enhance hydrolysis by up to 50-fold (164). Two other mutation have been identified in LcαE7 at one site, W251L or W251S, which are thought to reduce steric hindrance increasing the hydrolysis of OPs (63, 163, 164). So far, this mechanism has only been identified in other dipteran species, often closely

related to *L. cuprina*, such as *Musca domestica* and *Cochliomyia hominivorax* (137, 161, 165–167). It is thought that the fitness cost associated with this mutation has precluded its occurrence in other insect species (96, 168, 169). In *L. cuprina,* this cost is compensated by a mutation at another locus with an as of yet unknown function that rescues the fitness of *L. cuprina* and is widespread in *L. cuprina* populations (170).

2.1.5. Insect CBE expression

Our ability to understand how insecticide resistance CBEs function has been limited by the techniques that can be used to study them (30). The majority of effort has been spent in identifying insecticide resistance CBEs (10, 96, 144). This has been predominantly through transcriptome sequencing and RT-PCR (38, 128, 141, 171). This provides us with the sequence of CBE candidates and an association with insecticide resistance. While this is a critical first step, few studies have extended our understanding of these candidate CBEs through further characterization (109, 142, 157, 172). Those that have may rely on purification through complex separations using whole or partial insect body homogenates, which results in relatively small concentrations of the candidate CBE (4, 173–175). This enables basic biochemical characterization but is insufficient for structural characterization, which is key for both a molecular understanding of the insecticide resistance mechanism and for targeted-inhibitor design. For this, heterologous expression using *E. coli* expression systems is preferable due to its rapid growth and high expression of recombinant proteins (176–178).

2.1.6. *E. coli*-based expression of eukaryotic enzymes

E. coli expression systems have been utilized with relatively few insecticide resistance CBEs, to varying success (5, 160, 179, 180). Part of the difficulty relates to the differences between bacterial and eukaryotic protein translation (181). This includes all of the following present in eukaryotic cells but absent or reduced in prokaryotic cells: the ability to post-translationally modify proteins through the addition of sugar groups (glycosylation) or lipid groups (lipidation); and the presence of large suites of chaperones and organelle-based machinery to promote correct folding (181–183). In particular, the formation of structural disulfide bonds has been shown to be essential

for the function and stability of many proteins, especially those that are secreted (184). The lack of these systems can result in improper protein folding and the formation of insoluble aggregates or inclusion bodies (176, 185, 186). To combat this, specialized *E. coli* strains have been developed that have altered proteomes, eliminating detrimental proteins and adding beneficial ones, to promote the proper folding of eukaryotic proteins and thus expression in *E. coli* (182, 185, 187–189).

In this chapter, we utilized three *E. coli* strains to improve insect CBE expression: BL21(DE3) competent cells (NEB); Shuffle T7 Express Competent cells (NEB); and Origami B(DE3)pLysS competent cells (Novagen). BL21(DE3) cells are modified to be deficient in proteases (Lon and OmpT), resistant to phage T1 and to express a chromosomal copy of T7 RNA polymerase (T7 RNAP) allowing T7 expression, which is inducible by IPTG (190). The Origami B(DE3)pLysS contains the same modifications as BL21(DE3) cells in addition to: mutations in trxB (thioredoxin reductase) and gor (glutaredoxin reductase), which reduces their action and promotes disulfide bond formation; and the addition of T7 lysozyme, which suppresses T7 RNA polymerase prior to induction (190). Shuffle T7 Express cells have the largest number of modifications: the trxB and gor genes are deleted, promoting disulfide bond formation; a chromosomal copy of T7 RNAP is expressed; cells are deficient in Lon and OmpT; and a chromosomal copy of the disulfide bond isomerase, DsbC, is constitutively expressed, which promotes the correction of mis-oxidized proteins and acts as a chaperone for the proper folding of enzymes (189).

2.2. Preface

While CBE-mediated insecticide resistance is a growing problem, our efforts to combat it are limited by our molecular understanding of its mechanisms. In this chapter, I aim to further our understanding of insecticide resistance CBEs by identifying ideal candidates for large-scale expression and crystallization. I first conduct a broad literature search to find insecticide resistance CBEs that have been shown to be critical in resistance. These candidates are then subject to a number of computational analyses to determine their suitability for expression and crystallization. Of these, I use the most promising candidates in a range of expression trials to determine the ideal

conditions for large-scale expression. This is a critical first step in determining the structure of these insecticide resistance CBEs, which is essential to enhance our molecular understanding of insecticide resistance.

2.3. Materials and methods

2.3.1. Literature review and protein expression and crystallization prediction tools

Insecticide resistance CBE sequences were identified through a literature review. Signal peptides were detected for removal by the SignalP 4.1 server (191). Potential disulfide bonds were predicted using the DiANNA 1.1 web server (192). Potential solubility in *E. coli* was predicted using PROSO II (193). The protein crystallizability was predicted for each CBE using the XtalPred-RF web server (194).

2.3.2. Cloning

All insecticide resistance CBE amino acid sequences were obtained from UniProtKB or GenBank (195, 196). All DNA sequences were optimized for expression in *E. coli* and synthesized by Integrated DNA Technologies (IDT, USA) with N-terminal His-tags, TEV cleavage sites and N- and C-terminal sequences that overlap with the pETMCSIII vector to allow Gibson assembly (NEB) (197). DNA fragments were cloned into the pETMCSIII vector using Gibson assembly (197). Successful cloning was confirmed through DNA sequencing at the Biomolecular Resource Facility, Australian National University, Australia. Mutations, such as the addition of an S-tag, were introduced using overlapping primers and Gibson assembly (197). Plasmids for Hax42, Hax43 and Hax46 were provided by CSIRO Land and Water, Australia.

2.3.3. Protein expression

Proteins were expressed in three different *E. coli* strains: BL21(DE3) competent cells (NEB); Shuffle® T7 Express Competent cells (NEB); and Origami™ B(DE3)pLysS competent cells (Novagen) using the appropriate antibiotics. For small-scale expression each strain was grown in 10 ml lysogeny broth (LB) at varying

temperatures until induction at an OD600 of 0.4 - 0.8 using 0-1 mM IPTG. All cultures were then grown at room temperature for either 24 or 48 hours. In large-scale expression all conditions were maintained except 1 L of LB was used.

2.3.4. Protein lysis, separation and purification

For small-scale expression, cells were pelleted by centrifugation and resuspended in a BugBuster® solution (Merck, USA) containing resuspension buffer (20 mM HEPES pH 7.5, 150 mM NaCl) and turbonucease (Sigma, USA) and incubated at room temperature for 20 minutes. A sample of this 'whole cell' soluble and insoluble mixture was taken for protein separation before the lysed cells were pelleted and the soluble layer was removed for protein separation. In some cases, the insoluble pelleted fraction was then once again resuspended in BugBuster® solution and a sample of this insoluble fraction was taken for protein separation. Proteins were separated by SDS-PAGE using precast ExpressPlus 4 to 20% PAGE gels (GenScript) and stained using Coomassie brilliant blue (Sigma, USA) for visualization. For large-scale expression, cells were pelleted and resuspended in resuspension buffer before lysis by either sonication or French press. Cell debris was pelleted by centrifugation and the soluble fraction was applied to a HisTrap FF column (GE Healthcare). Bound protein was eluted using buffer B (20 mM HEPES pH 7.5, 150 mM NaCl, 300 mM imidazole).

2.4. Results

2.4.1. Identification of candidate genes

To determine ideal insecticide resistance CBE genes to target for expression and crystallization I conducted a comprehensive search in the literature. While there are many studies that identify a potential link between resistance and CBE expression, there are far fewer that determine the sequence of these CBEs and even fewer that directly analyze them (10). From those studies that have empirically analyzed CBEs, I chose to focus on insecticide resistance CBEs that have already been well studied, such as those from *C. quinquefasciatus* and *Myzus persicae* (30). I also focused on CBEs from pests of significant impact to humanity such as *Helicoverpa armigera*, one

of the most important caterpillar pests of cotton in many countries (198). In total, I identified 23 insecticide resistance CBE candidates for further analysis and potential expression and crystallization using *E. coli*. Within this set of proteins, there was a decreasing level of amino acid identity to the best studied catalytic insecticide resistance CBE, LcαE7 from *L. cuprina* (**Table 2.1**).

Table 2.1. A list of insecticide resistance CBEs identified for future analysis.

Organism	Order	Uniprot ID	Name	Seq ID with LcαE7 (%)	Reference
Bactrocera dorsalis	Diptera	A0A0H4DA18_BACDO	BdB1	58	(174)
Ceratitis capitata	Diptera	B7SB38_CERCA	CcαE7	57	(199)
C. quinquefasciatus	Diptera	P92025_CULQU	Cqestα2[1]	41	(200)
Aedes aegypti	Diptera	Q16T49_AEDAE	AaB1	41	(201)
A. gambiae	Diptera	Q7PPB0_ANOGA	AgB2	39	(202)
C. quinquefasciatus	Diptera	Q23734_CULQU	Cqestβ2[1]	38	(31)
Aphis gossypii	Hemiptera	O76177_APHGO	CarE-YS3	37	(203)
Locusta migratoria	Orthoptera	T1VXB0_LOCMI	CesA4	37	(204)
Anisopteromalus calandrae	Hymenoptera	O61726_9HYME	Ac-CCE	36	(150)
Bemisia tabaci	Hemiptera	B3SST0_BEMTA	COE1	35	(149)
Cydia pomonella	Lepidoptera	A0A0C4JZC5_CYDPO	CpCE-1	35	(159)
A. aegypti	Diptera	Q17B28_AEDAE	cceae6a	35	(201)
A. aegypti	Diptera	Q17B31_AEDAE	cceae3a	34	(201)
Nilaparvata lugens	Lepidoptera	Q9GQ01_NILLU	Nl-EST1	34	(205)
H. armigera	Lepidoptera	(unpublished)[a]	Hax42	33	(unpublished)
M. persicae	Hemiptera	ESTF_MYZPE	MpFE4	32	(206)
M. persicae	Hemiptera	ESTE_MYZPE	MpE4	31	(179)
H. armigera	Lepidoptera	A0A0U3BZB1_HELAM	Hax001D	31	(180)
Leptinotarsa decemlineata	Coleoptera	A0A0A7EP32_LEPDE	Km220566	31	(40)
H. armigera	Lepidoptera	(unpublished)[a]	Hax46	29	(unpublished)
H. armigera	Lepidoptera	(unpublished)[a]	Hax43	28	(unpublished)
Rhipicephalus microplus	Ixodida	Q9U6M8_RHIMP	RmEST9	28	(207)

[a] Enzymes were provided by Dr. John Oakeshott, CSIRO Land and Water, Australia.

The CBE Cqestβ2[1] from *C. quinquefasciatus* was of particular interest due to three key factors: (i) it has been suggested to act through sequestration; (ii) a closely related CBE, Cqestβ1, has been expressed in *E. coli*; and (iii) the enzyme was shown to be unaffected by the equivalent to the G137D OP hydrolase gain-in-function mutation found in LcαE7 (31, 33, 173, 208).

2.4.2. Computational predictions of CBE expression and crystallization

To determine the likelihood of expression in *E. coli* and propensity for crystallization, I utilized a number of computational prediction tools. PROSO II is a protein solubility prediction tool that uses a machine-learning-based model that has been trained on 82000 proteins and is suggested to have an accuracy of 75.4% (193). This tool found that 12 of the 23 candidates were likely to have soluble expression in *E. coli* (**Table 2.2**). These predictions largely agree with the limited previous heterologous

expression data with the exception of MpE4, which is predicted to be insoluble (159, 179, 180, 208). As disulfide bonds are fairly common in insect enzymes and can create issues with expression in *E. coli*, I utilized the disulfide bond prediction tool, DiANNA (192). This determined that 20 of the 23 candidates were likely to have disulfide bonds (**Table 2.2**). Due to the presence of specialized *E. coli* strains for the expression of proteins with disulfide bonds, this was not used as a criterion to disregard a candidate for expression (177). To determine the likelihood of crystallization I used the XtalPred web server (194). This uses two methods to predict "crystallizability": the Expert Pool method (EP class); and the Random Forest Classifier (RF class) (194). EP class uses eight protein features to generate a score that bins the candidate into one of five crystallization classes where 1 is the most likely to crystallize (194, 209). RF class uses additional features to generate the score and uses eleven classes where 1 is the most likely to crystallize (194, 210). Five candidates were found to be in the top two EP class and eleven in the top three RF class (**Table 2.2**).

Table 2.2. A summary of the results from a range of computational prediction tools used to determine the suitability for each CBE candidate for further research.

Candidate	PROSO II[a]	DiANNA	XtalPred EP Class[a] (1-5)	XtalPred RF Class[a] (1-11)	Previous heterologous expression
Cqestβ2[1]	soluble, 0.734	unlikely	3	1	*E. coli* (208)
Cqestα2[1]	soluble, 0.631	likely	2	3	none
AgB2	insoluble, 0.596	unlikely	4	1	none
BdB1	soluble, 0.630	likely	2	5	Baculovirus (174)
AaB1	soluble, 0.644	likely	3	1	none
CcαE7	soluble, 0.743	likely	4	11	none
cceae3a	insoluble, 0.557	likely	4	4	none
cceae6a	insoluble, 0.533	likely	3	4	none
Hax001D*	soluble, 0.639	likely	4	3	*E. coli* (180)
Hax42*	insoluble, 0.583	likely	4	2	none
Hax43	soluble, 0.697	likely	4	3	none
Hax46	soluble, 0.736	likely	3	2	none
Nl-EST1*	soluble, 0.712	likely	3	3	none
CpCE-1*	soluble, 0.674	likely	3	4	*E. coli* (159)
MpE4*	insoluble, 0.469	likely	3	6	*E. coli* (179)
MpFE4*	insoluble, 0.504	likely	2	4	none
CarE-YS3	insoluble, 0.589	likely	2	3	none
COE1	soluble, 0.653	likely	4	3	none
Km220566	soluble, 0.702	likely	3	5	none
RmEST9*	insoluble, 0.412	likely	4	5	none
CesA4	insoluble, 0.507	likely	4	7	none
Ac-CCE	insoluble, 0.357	unlikely	2	11	none

[a] Enzymes labelled with an * after their name possessed signal peptides that were removed for solubility and crystallization predictions. Enzymes that were selected for expression trials are shown in bold.

In total, thirteen candidates were selected for expression trials based on the above prediction tools and evidence of heterologous expression (**Table 2.2**). I also aimed to

include candidates from a range of insect orders taking the significance of each insect pest into consideration.

2.4.3. Expression trials

Each selected insecticide resistance CBE candidate was optimized and cloned into the pETMCSIII vector for expression in *E. coli*. Previous studies had confirmed expression of the *H. armigera* CBE, Hax001D in BL21(DE3) cells using a combination of a His-tag and a solubility tag (S-tag), which was incorporated in its sequence (180). Candidates were first expressed in BL21(DE3) competent cells and if no clear soluble expression was detected they were expressed in both Origami B(DE3)pLysS competent cells and Shuffle T7 Express Competent cells. I incorporated an S-tag to a number of candidates without clear soluble expression to test its effect. The results of each candidate are summarized below:

Expression of the C. quinquefasciatus CBE, Cqestβ2[1]

While there is no soluble expression present for Cqestβ2[1] in BL21(DE3) cells without induction, a large amount is present upon induction with IPTG (**Figure 2.1**).

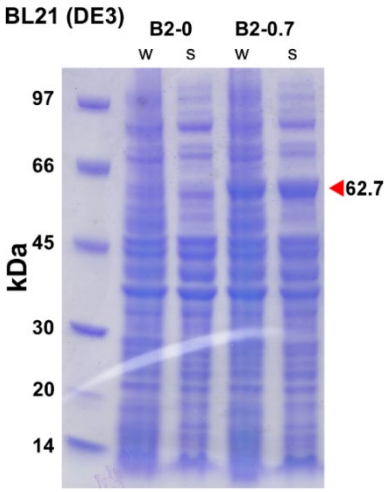


Figure 2.1. Expression trial of *Cq*estβ2[1] using BL21(DE3) *E. coli* competent cells and varied IPTG concentrations for induction: “B2-0” indicates expression with no induction; “B2-0.7” indicates induction with 0.7 mM IPTG; “w” refers to the whole cell fraction; and “s” refers to the soluble fraction. Target protein size is indicated by the red arrow and labelled.

Expression of the C. quinquefasciatus CBE, Cqestα2[1]

There is no clear insoluble or soluble expression of Cqestα2[1] or Cqestα2[1] with an S-tag in BL21(DE3) cells (**Figure 2.2**) or in Origami B(DE3)pLysS cells (**Figure 2.3**). There is a small amount of insoluble expression of both Cqestα2[1] and Cqestα2[1] with an S-tag in Shuffle T7 Express cells (**Figure 2.4**).

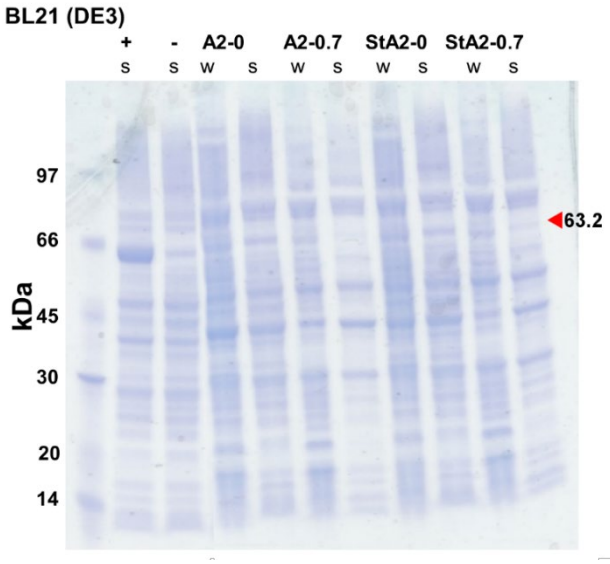


Figure 2.2. Expression trial of Cqestα2[1] and S-tagged Cqestα2[1] using BL21(DE3) *E. coli* competent cells without induction or with induction using 0.7 mM IPTG: "+" refers to a positive control using LcαE7; "-" refers to a negative control using empty vector; "A2-0" refers to Cqestα2[1] expression without induction and "A2-0.7" refers to its expression with induction; "StA2-0" refers to S-tagged Cqestα2[1] without induction and "StA2-0.7" refers to its expression with induction; "w" refers to the whole cell fraction; and "s" refers to the soluble fraction. Two irrelevant protein samples on the same gel were removed between samples "–" and "A2-0, w". Target protein size is indicated by the red arrow and labelled.

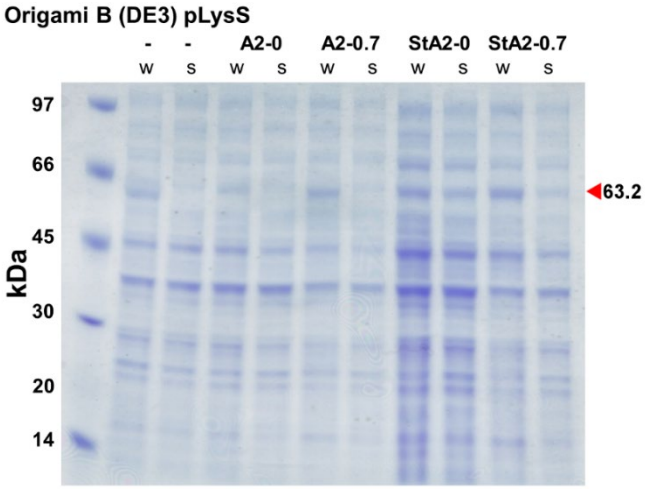


Figure 2.3. Expression trial of Cqestα2[1] and S-tagged Cqestα2[1] using Origami B(DE3)pLysS *E. coli* competent cells without induction or with induction using 0.7 mM IPTG: "-" refers to a negative control using empty vector; "A2-0" refers to Cqestα2[1] expression without induction and "A2-0.7" refers to its expression with induction; "StA2-0" refers to S-tagged Cqestα2[1] without induction and "StA2-0.7" refers to its expression with induction; "w" refers to the whole cell fraction; and "s" refers to the soluble fraction. Target protein size is indicated by the red arrow and labelled.

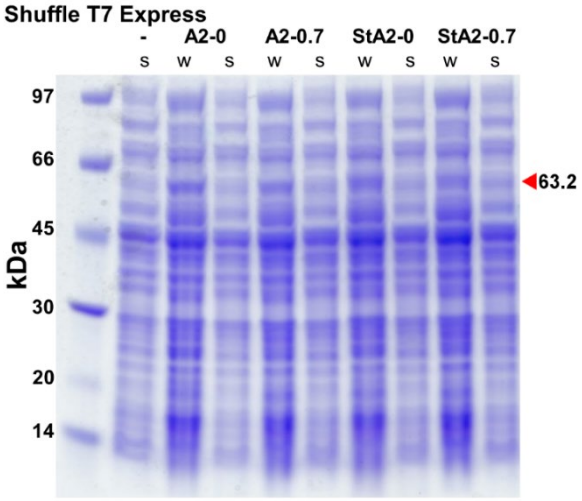


Figure 2.4. Expression trial of Cqestα2[1] and S-tagged Cqestα2[1] using Shuffle T7 Express *E. coli* competent cells without induction or with induction using 0.7 mM IPTG: "-" refers to a negative control using empty vector; "A2-0" refers to Cqestα2[1] expression without induction and "A2-0.7" refers to its expression with induction; "StA2-0" refers to S-tagged Cqestα2[1] without induction and "StA2-0.7" refers to its expression with induction; "w" refers to the whole cell fraction; and "s" refers to the soluble fraction. Target protein size is indicated by the red arrow and labelled.

Expression of the A. aedes CBE, AaB1

There is a moderate amount of soluble expression of AaB1 without induction and a large amount with induction in BL21(DE3) cells (**Figure 2.5**).

Expression of the A. gambiae CBE, AgB2

There is a moderate amount of soluble expression of AgB2 without induction and a large amount with induction in BL21(DE3) cells (**Figure 2.5**).

Expression of the B. dorsalis CBE, BdB1

There is no clear soluble or insoluble expression of BdB1 without induction but strong insoluble and low levels of soluble expression with induction in BL21(DE3) cells (**Figure 2.5**).

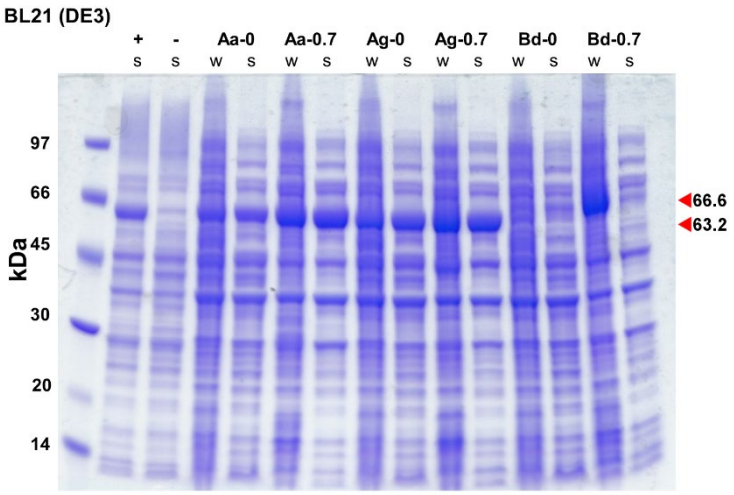


Figure 2.5. Expression trial of AaB1, AgB2 and BdB1 using BL21(DE3) *E. coli* competent cells without induction or with induction using 0.7 mM IPTG: "+" refers to a positive control using LcαE7; "-" refers to a negative control using empty vector; "Aa-0" refers to AaB1 expression without induction and "Aa-0.7" refers to its expression with induction; "Ag-0" refers to AgB2 expression without induction and "Ag-0.7" refers to its expression with induction; "Bd-0" refers to BdB1 expression without induction and "Bd-0.7" refers to its expression with induction; "w" refers to the whole cell fraction; and "s" refers to the soluble fraction. Target protein size is indicated by the red arrow and labelled: 66.6 kDa for BdB1 and 63.2 kDa for AaB1. AgB2 is 62.1kDa and overlaps AaB1 so is not shown.

Expression of the C. capitata CBE, CcαE7

In BL21(DE3) cells there is no clear expression of CcαE7 without induction, however, with induction there are moderate levels of insoluble and soluble expression (**Figure 2.6**).

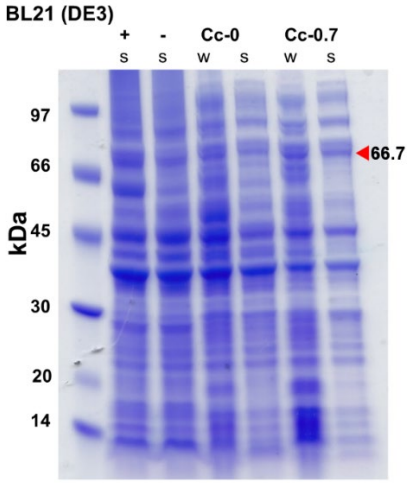


Figure 2.6. Expression trial of CcαE7 using BL21(DE3) *E. coli* competent cells and varied IPTG concentrations for induction: "+" refers to a positive control using LcαE7; "-" refers to a negative control using empty vector; "Cc-0" refers to CcαE7 expression without induction; "Cc-0.7" refers to its expression with induction using 0.7 mM IPTG; "w" refers to the whole cell fraction; and "s" refers to the soluble fraction. Target protein size is indicated by the red arrow and labelled.

Expression of the M. persicae CBE, MpE4

There is no noticeable difference between the soluble expression of empty vector and MpE4 with or without an S-tag in BL21(DE3) cells, indicating no soluble expression (**Figure 2.7**). There is strong insoluble expression of both MpE4 and S-tagged MpE4 in BL21(DE3) cells (**Figure 2.7**). Unfortunately, there is a protein of a similar size to MpE4 in empty vector expression of Origami B(DE3)pLysS cells making it harder to interpret the results (**Figure 2.8, 2.9**). It is still unlikely that there is any soluble expression of MpE4 with or without an S-tag in either Origami B(DE3)pLysS or Shuffle T7 Express cells (**Figure 2.8, 2.9**). There is no insoluble expression with MpE4 in either Origami B(DE3)pLysS or Shuffle T7 Express cells but low levels with S-tagged MpE4 in Origami B(DE3)pLysS cells and moderate levels in Shuffle T7 Express cells with and without induction (**Figure 2.8, 2.9**). In general, the addition of an S-tag seems to increase the levels of insoluble expression.

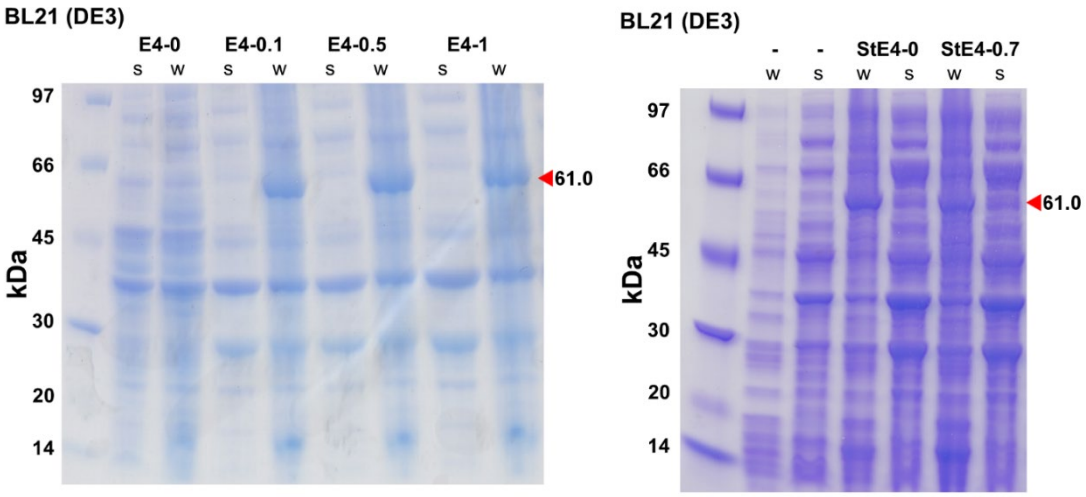


Figure 2.7. Expression trials of MpE4 and S-tagged MpE4 using BL21(DE3) *E. coli* competent cells and varied IPTG concentrations for induction: "E4-0" refers to MpE4 expression without induction, "E4-0.1", "E4-0.5" and "E4-1" refer to its expression with induction using 0.1 mM, 0.5 mM and 1 mM IPTG, respectively; "-" refers to a negative control using empty vector; "StE4-0" refers to S-tagged MpE4 expression without induction and "StE4-0.7" refers to its expression with induction using 0.7 mM IPTG; "w" refers to the whole cell fraction; and "s" refers to the soluble fraction. Target protein size is indicated by the red arrow and labelled.

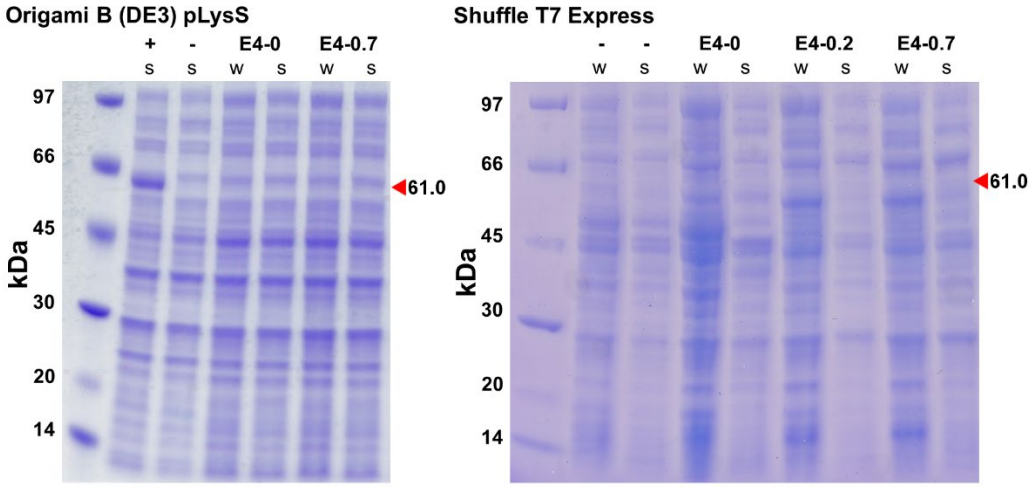


Figure 2.8. Expression trials of MpE4 using Origami B(DE3)pLysS and Shuffle T7 Express *E. coli* competent cells and varied IPTG concentrations for induction: "+" refers to a positive control using Esterase 6-1 from *D. melanogaster* (DmEST6-1) (24); "-" refers to a negative control using empty vector; "E4-0" refers to MpE4 expression without induction, "E4-0.2" and "E4-0.7" refer to its expression with induction using 0.2 mM and 0.7 mM IPTG, respectively; "w" refers to the whole cell fraction; and "s" refers to the soluble fraction. Target protein size is indicated by the red arrow and labelled.

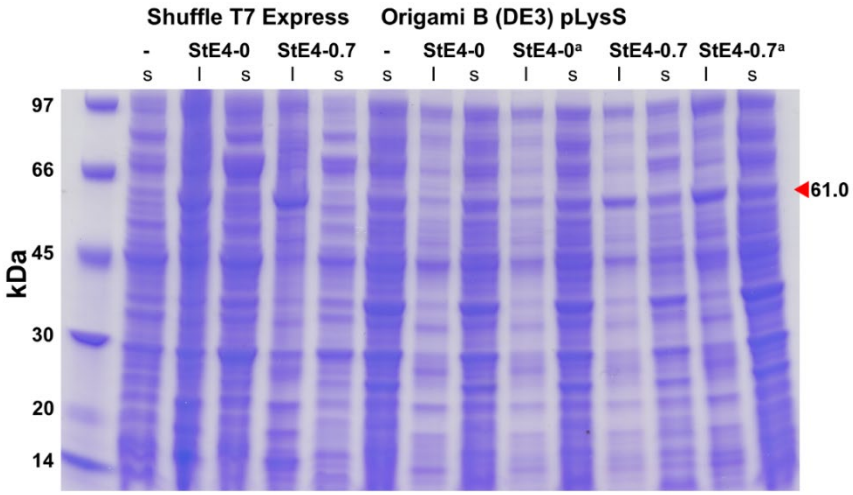


Figure 2.9. Expression trial of S-tagged MpE4 using Shuffle T7 Express *E. coli* competent cells, Origami B(DE3)pLysS *E. coli* competent cells and varied IPTG concentrations for induction: "-" refers to a negative control using empty vector; "StE4-0" refers to S-tagged MpE4 expression without induction and "StE4-0.7" refers to its expression with induction using 0.7 mM IPTG; "I" refers to the insoluble fraction; and "s" refers to the soluble fraction. "ª" indicates that expression was extended for 48 hours after induction. Target protein size is indicated by the red arrow and labelled.

Expression of the M. persicae CBE, MpFE4

Similar to MpE4, there is a protein of a similar size to MpFE4 in empty vector expression in BL21(DE3) and Origami B(DE3)pLysS cells making it harder to interpret the results (**Figure 2.10 – 2.12**). However, it still appears unlikely that there is soluble expression of MpFE4, with or without an S-tag, in either Bl21(DE3) or Origami B(DE3)pLysS cells (**Figure 2.10 – 2.12**). There are large amounts of insoluble protein found in BL21(DE3) cells with induction (**Figure 2.10**) and S-tagged MpFE4 in Shuffle T7 Express with or without induction (**Figure 2.12**). A moderate amount of insoluble expression is clear in S-tagged MpFE4 samples with or without induction in BL21(DE3) cells (**Figure 2.12**) and with MpFE4 in Shuffle T7 Express cells with induction (**Figure 2.11**). As with MpE4, the addition of an S-tag seems to increase the levels of insoluble expression in some conditions.

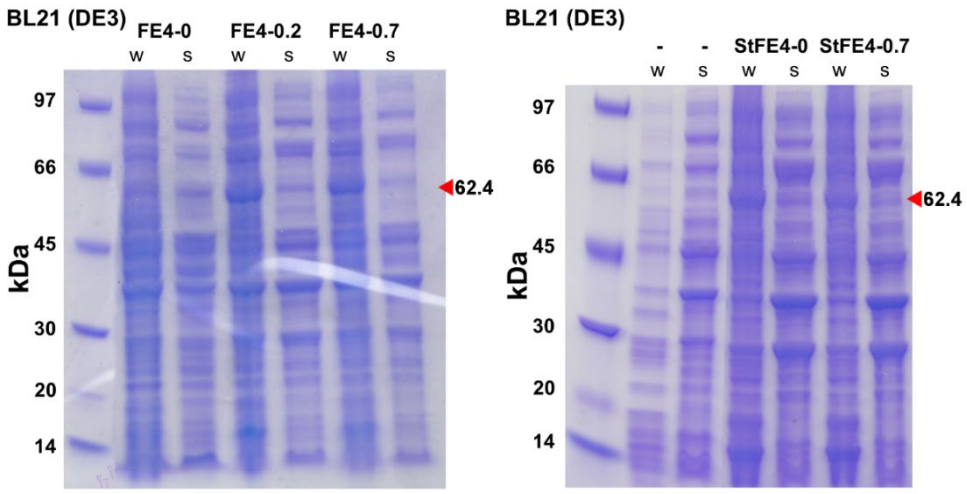


Figure 2.10. An expression trial of MpFE4 and an expression trial of S-tagged MpFE4 using BL21(DE3) *E. coli* competent cells and varied IPTG concentrations for induction: "-" refers to negative controls using empty vector; "FE4-0" refers to MpFE4 expression without induction; "FE4-0.2" and "FE4-0.7" refer to MpFE4 expression with induction using 0.2 mM and 0.7 mM IPTG, respectively; "StFE4-0" refers to S-tagged MpFE4 expression without induction and "StFE4-0.7" refers to its expression with induction using 0.7 mM IPTG; "w" refers to the whole cell fraction; and "s" refers to the soluble fraction. Four irrelevant protein samples on both gels were removed between ladder and sample "FE4-0, w" and samples "-, s" and "StFE4-0, w", respectively. Target protein size is indicated by the red arrow and labelled.

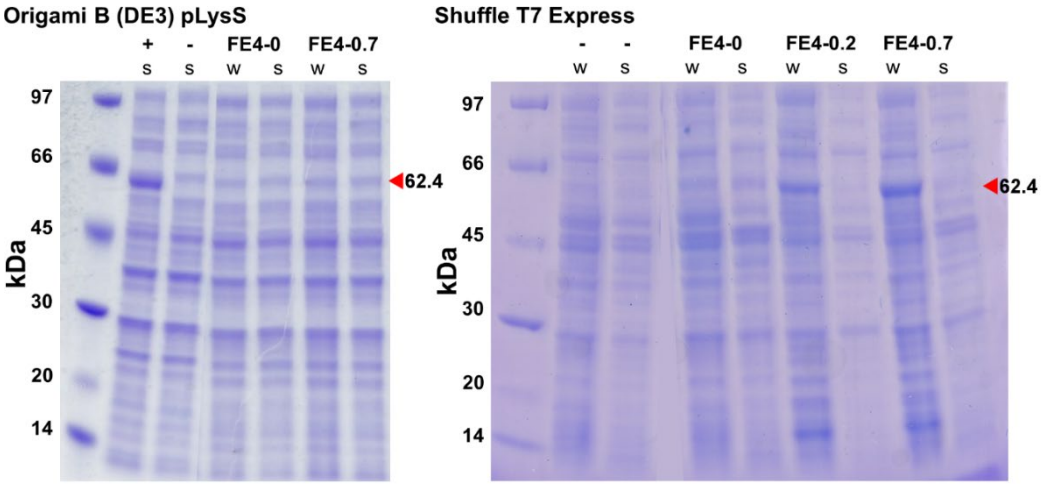


Figure 2.11. Expression trials of MpFE4 using Origami B(DE3)pLysS and Shuffle T7 Express *E. coli* competent cells and varied IPTG concentrations for induction: "+" refers to a positive control using DmEST6-1; "-" refers to a negative control using empty vector; "FE4-0" refers to MpFE4 expression without induction and "FE4-0.2" and "FE4-0.7" refer to its expression with induction using 0.2 mM and 0.7 mM IPTG, respectively; "w" refers to the whole cell fraction; and "s" refers to the soluble fraction. Four irrelevant protein samples were removed from both gels between samples "-, s" and "FE4-0, w". Target protein size is indicated by the red arrow and labelled.

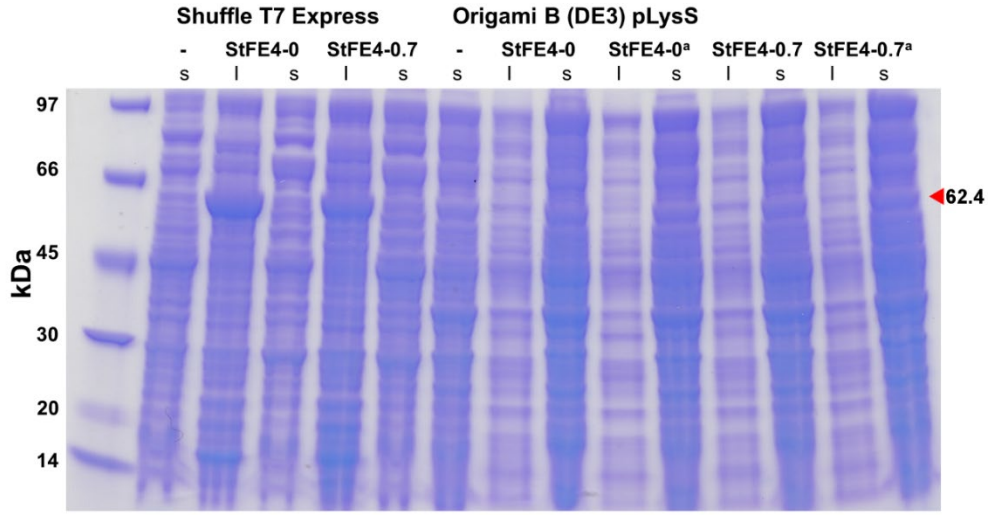


Figure 2.12. Expression trial of S-tagged MpFE4 using Shuffle T7 Express *E. coli* competent cells, Origami B(DE3)pLysS *E. coli* competent cells and varied IPTG concentrations for induction: "-" refers to a negative control using empty vector; "StFE4-0" refers to S-tagged MpFE4 expression without induction and "StFE4-0.7" refers to its expression with induction using 0.7 mM IPTG; "I" refers to the insoluble fraction; and "s" refers to the soluble fraction. "ª" indicates that expression was extended for 48 hours after induction. Target protein size is indicated by the red arrow and labelled.

Expression of the L. decemlineata CBE, Km220566

In BL21(DE3) cells there appears to be a low level of insoluble expression of both induced and uninduced Km220566, however, there is clearly no soluble expression (**Figure 2.13**). In both Origami B(DE3)pLysS and Shuffle T7 Express cells there is no difference in either insoluble or soluble expression between Km220566 and empty vector, indicating no expression (**Figure 2.13**).

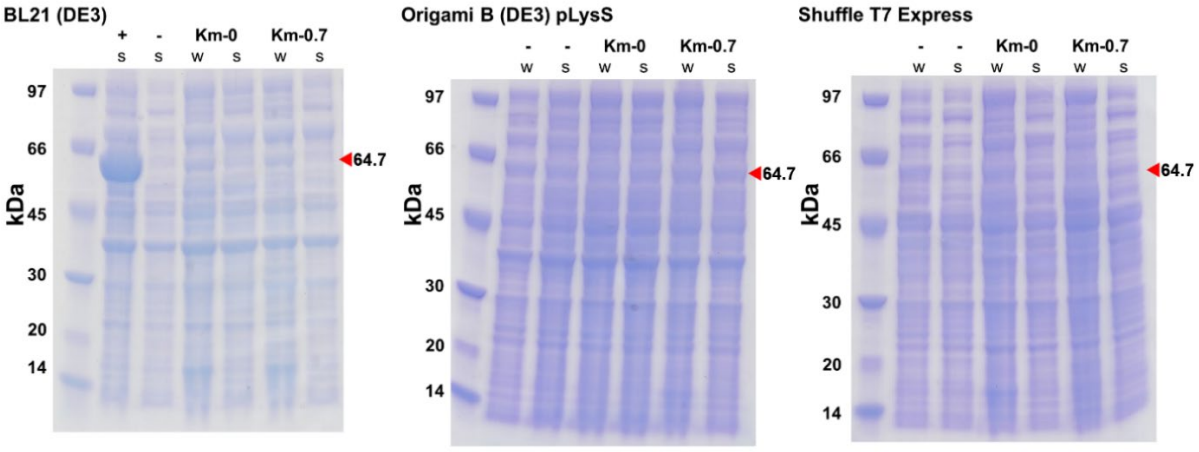


Figure 2.13. Expression trials of Km220566 using BL21(DE3), Origami B(DE3)pLysS and Shuffle T7 Express *E. coli* competent cells with varied IPTG concentrations for induction: "+" refers to a positive control using LcαE7; "-" refers to a negative control using empty vector; "Km-0" refers to Km220566 expression without induction and "Km-0.7" refers to its expression with induction using 0.7 mM IPTG; "w" refers to the whole cell fraction; and "s" refers to the soluble fraction. Target protein size is indicated by the red arrow and labelled.

Expression of the H. armigera CBE, Hax001D

A comparison with the expression of empty vector suggests that there is no soluble or insoluble expression of Hax001D in either Origami B(DE3)pLysS or BL21(DE3) cells, with or without induction (**Figure 2.14**). While there may be a low level of soluble expression in Shuffle T7 Express cells with induction, it is more likely that the sample was loaded at a higher concentration than the negative control, due to the similarly increased intensity of other bands in the sample over the control. This appears to also be an issue with the whole cell samples of both induced and uninduced Hax001D in Shuffle T7 Express cells (**Figure 2.14**). Thus, it is also unlikely there is either soluble or insoluble expression of Hax001D in Shuffle T7 Express cells.

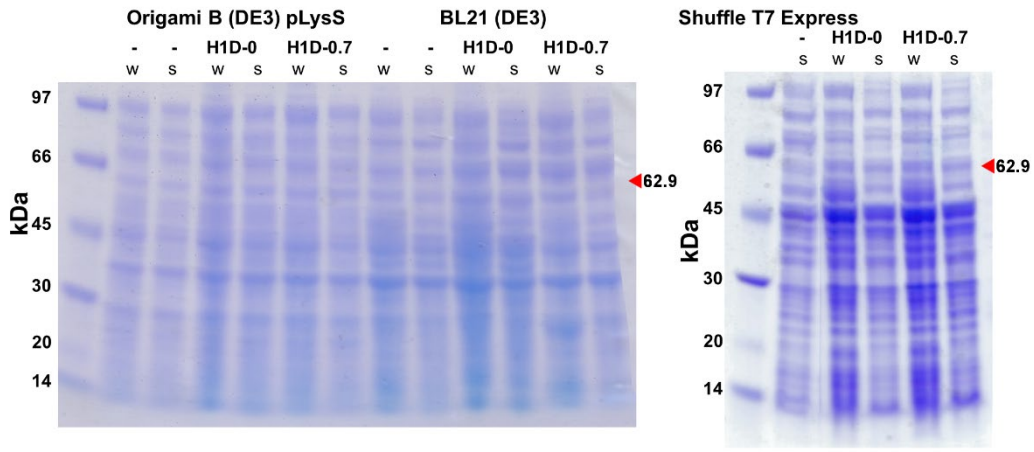


Figure 2.14. Expression trials of Hax001D using BL21(DE3), Origami B(DE3)pLysS and Shuffle T7 Express *E. coli* competent cells with varied IPTG concentrations for induction: "-" refers to a negative control using empty vector; "H1D-0" refers to Hax001D expression without induction and "H1D-0.7" refers to its expression with induction using 0.7 mM IPTG; "w" refers to the whole cell fraction; and "s" refers to the soluble fraction. Target protein size is indicated by the red arrow and labelled.

Expression of the H. armigera CBE, Hax42

In BL21(DE3) cells there is no clear difference between the soluble expression of Hax42 and empty vector, indicating no soluble expression of Hax42 (**Figure 2.15**). However, there is strong insoluble expression in induced BL21(DE3) cells (**Figure 2.15**). Similarly, induction resulted in the production of moderate levels of insoluble expression with Origami B(DE3)pLysS cells but no clear soluble expression (**Figure 2.16**). While there is no soluble expression in Shuffle T7 Express cells, there is an increase in the levels of insoluble expression in the uninduced sample (**Figure 2.17**).

Expression of the H. armigera CBE, Hax43

There is no clear soluble expression of Hax43 in BL21(DE3) cells with or without induction (**Figure 2.15**). There is, however, a large amount of insoluble expression in induced BL21(DE3) cells, similar to Hax42 (**Figure 2.15**). In Origami B(DE3)pLysS cells there is neither soluble nor insoluble expression (**Figure 2.16**). In Shuffle T7 Express cells there does appear to be a moderate level of insoluble expression of Hax43 with and without induction, however, there is no soluble expression (**Figure 2.17**).

Expression of the H. armigera CBE, Hax46

There is no clear difference between the empty vector soluble expression and soluble or insoluble expression of Hax46, with or without induction, in BL21(DE3) cells (**Figure 2.15**) Origami B(DE3) cells (**Figure 2.16**) or Shuffle T7 Express cells (**Figure 2.17**). This suggests that there is neither soluble nor insoluble expression for Hax46 in any of the conditions tested.

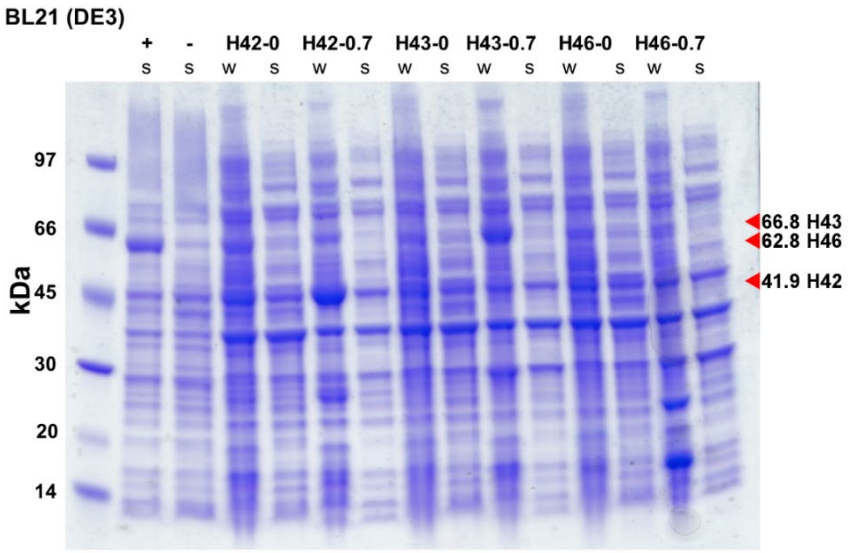


Figure 2.15. Expression trial of Hax42, Hax43 and Hax46 using BL21(DE3) *E. coli* competent cells without induction or with induction using 0.7 mM IPTG: "+" refers to a positive control using LcαE7; "-" refers to a negative control using empty vector; "H42-0" refers to Hax42 expression without induction and "H42-0.7" refers to its expression with induction; "H43-0" refers to Hax43 expression without induction and "H43-0.7" refers to its expression with induction; "H46-0" refers to Hax46 expression without induction and "H46-0.7" refers to its expression with induction; "w" refers to the whole cell fraction; and "s" refers to the soluble fraction. Target protein sizes are indicated by the red arrow and labelled by size and name.

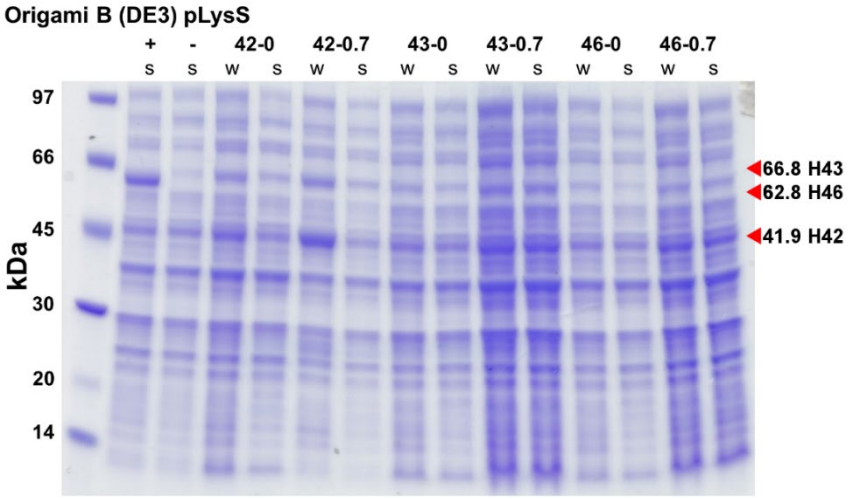


Figure 2.16. Expression trial of Hax42, Hax43 and Hax46 using Origami B(DE3)pLysS *E. coli* competent cells without induction or with induction using 0.7 mM IPTG: "+" refers to a positive control using DmEST6-1; "-" refers to a negative control using empty vector; "H42-0" refers to Hax42 expression without induction and "H42-0.7" refers to its expression with induction; "H43-0" refers to Hax43 expression without induction and "H43-0.7" refers to its expression with induction; "H46-0" refers to Hax46 expression without induction and "H46-0.7" refers to its expression with induction; "w" refers to the whole cell fraction; and "s" refers to the soluble fraction. Target protein sizes are indicated by the red arrow and labelled by size and name.

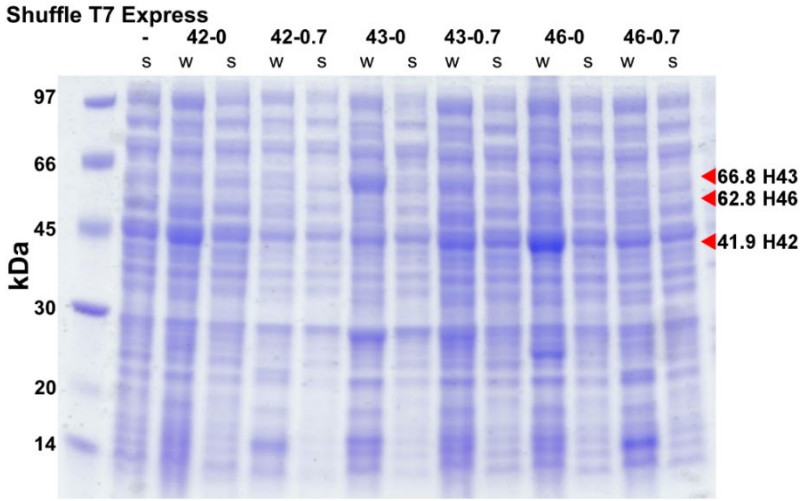


Figure 2.17. Expression trial of Hax42, Hax43 and Hax46 using Shuffle T7 Express *E. coli* competent cells without induction or with induction using 0.7 mM IPTG: "-" refers to a negative control using empty vector; "H42-0" refers to Hax42 expression without induction and "H42-0.7" refers to its expression with induction; "H43-0" refers to Hax43 expression without induction and "H43-0.7" refers to its expression with induction; "H46-0" refers to Hax46 expression without induction and "H46-0.7" refers to its expression with induction; "w" refers to the whole cell fraction; and "s" refers to the soluble fraction. Target protein sizes are indicated by the red arrow and labelled by size and name.

In summary, the expression trials identified five insecticide resistance CBEs that display some level of soluble expression (BdB1, CcαE7, Cqestβ2[1], AaB1 and AgB2) and three with expression ideal for utilization in crystallization trials (Cqestβ2[1], AaB1 and AgB2) (**Table 2.3**).

Table 2.3. A summary of the results for CBE expression. +IPTG indicates the addition of 0.7 mM IPTG to cell cultures at OD600 0.4 – 0.8 whereas -IPTG indicates no addition. Strength of protein bands were decided qualitatively and are indicated by intensity of color where green indicates soluble expression and blue indicates insoluble expression.

	Solubility fraction	BL21(DE3) -IPTG	BL21(DE3) +IPTG	Origami B(DE3)pLysS -IPTG	Origami B(DE3)pLysS +IPTG	Shuffle T7 Express -IPTG	Shuffle T7 Express +IPTG
Cqestβ2[1]	insoluble	-	low				
	soluble	-	strong				
AaB1	insoluble	low	low				
	soluble	medium	strong				
AgB2	insoluble	low	low				
	soluble	medium	strong				
BdB1	insoluble	-	strong				
	soluble	-	low				
CcαE7	insoluble	-	medium				
	soluble	-	medium				
Cqestα2[1]	insoluble	-	-	-	-	low	low
	soluble	-	-	-	-	-	-
Cqestα2[1] w S-tag	insoluble	-	-	low	low	medium	medium
	soluble	-	-	-	-	-	-
MpE4	insoluble	-	strong	-	-	-	-
	soluble	-	-	-	-	-	-
MpE4 w S-tag	insoluble	strong	medium	low	low	medium	medium
	soluble	-	-	-	-	-	-
MpFE4	insoluble	-	strong	-	-	-	medium
	soluble	-	-	-	-	-	-
MpFE4 w S-tag	insoluble	medium	medium	-	-	strong	strong
	soluble	-	-	-	-	-	-
Km220566	insoluble	low	low	-	-	-	-
	soluble	-	-	-	-	-	-
Hax001D	insoluble	-	-	-	-	-	-
	soluble	-	-	-	-	-	-
Hax42	insoluble	-	strong	-	medium	medium	-
	soluble	-	-	-	-	-	-
Hax43	insoluble	-	strong	-	-	medium	medium
	soluble	-	-	-	-	-	-
Hax46	insoluble	-	-	-	-	-	-
	soluble	-	-	-	-	-	-

2.4.4. Large-scale expression trials of insoluble proteins

There were four insecticide resistance CBE candidates that showed strong insoluble expression but no soluble expression: MpE4, MpFE4, Hax42 and Hax43. For each candidate I attempted a large-scale expression using 1 L of LB in the same growth conditions. As different protein extraction techniques can affect the amount of soluble protein, I tried both French press and sonication techniques opposed to BugBuster, which was used for the small-scale expressions (211). Purification using affinity chromatography did not yield any soluble protein (**data not shown**).

2.5. Discussion

To better understand the biochemistry and structure of insecticide resistance CBEs we must first be able to express them. To this end, I identified a number of CBE candidates, used computational tools to predict their viability for expression and crystallization and conducted a range of expression trials with thirteen of them. Of these, I identified five candidates that demonstrate some level of expression: BdB1, CcαE7, Cqestβ2[1], AaB1 and AgB2. Each of these enzymes had soluble expression with BL21(DE3) *E. coli* competent cells. Of all the candidates tested, these five had some of the highest sequence identities with LcαE7 (from 38 – 58%), which is also known to express in BL21(DE3) cells (**Table 2.1**) (5). This similarity may have been involved in their successful expression. None of them required the removal of a signal peptide, which may have also played a role in their successful expression. While both these enzymes and LcαE7 are from organisms in the insect order Diptera, the lack of expression with the dipteran Cqestα2[1] suggests other factors may be key to their expression. Interestingly, the predicted presence of disulfide bonds did not prevent the expression of BdB1, AaB1 or CcαE7 in BL21(DE3) cells, which are not designed to promote disulfide bond formation. The PROSO II server correctly predicted the soluble expression of four of the candidates Cqestβ2[1], BdB1, CcαE7 and AaB1, which had scores ranging from 0.630 – 0.743 (**Table 2.2**). In total, the PROSO II server correctly predicted the solubility of eight of the thirteen candidates suggesting it is a reasonable, but far from definitive, tool for solubility estimation.

Previous studies have shown successful heterologous expression for both MpE4 and Hax001D using *E. coli* expression systems. However, I was unable to detect any soluble expression for either in my expression trials (**Table 2.3**) (179, 180). There are a number of differences that may have affected each enzyme's expression in my trials. One key difference is the vector used for expression: for the expression of Hax001D they tried a number of vectors (pE1, pET32a and pET30a) but only found decent expression and purity with pET30a (180); and for the expression of MpE4 they used the pET28b vector (179). While pET28b, pET30a and pETMCSIII, which I used, are all high expression, T7 promoter vectors from the pET series, as shown with Hax001D, different vectors, even similar ones, can have a large impact on expression (178, 180). The growth medium used with Hax001D was also supplemented with casein hydrolysate, which may have also improved its expression (180).

The majority of insecticide resistance CBEs that I tested were found to only have insoluble expression (**Table 2.3**). Apart from Cqestα2[1] these CBEs all share a sequence identity of less than 34% with LcαE7 and were all predicted to possess disulfide bonds (**Table 2.1, Table 2.2**). While this does not preclude expression in *E. coli* it can make it more difficult. It is also possible that these CBEs require post translational modifications, such as glycosylation, acylation, phosphorylation and acetylation, or folding machinery not present in the *E. coli* strains I used to generate soluble protein (181, 188). This is a common feature of eukaryotic enzymes that complicates expression in *E. coli* (181, 188, 212). Interestingly, while the addition of an S-tag to enzymes did not promote greater soluble expression, it did result in a greater production of insoluble protein in the same conditions. This suggests it may have increased the incorrectly folding proteins resistance to degradation and thus the propensity for protein aggregation into inclusion bodies.

2.6. Further research

The identification of soluble expression in a number of the insecticide resistance CBEs enabled the large-scale expression and crystallization of Cqestβ2[1] described in the following chapter. It also enabled other members of Assoc. Prof. Colin Jackson's research group to use BdB1, CcαE7, AgB2, AaB1 and Cqestβ2[1] in an investigation of

the structural limitations in insecticide resistance CBEs related to the G137D mutation in LcαE7 and to determine the crystal structure of AgB2 (unpublished).

The significant occurrence of insoluble expression with the insecticide resistance CBEs suggests that alternate strategies should be utilized to obtain soluble expression. One strategy that was not used in this study would be to clone each insecticide resistance CBE into a low copy number vector, such as pWSK29 and pWKS30, and to utilize vectors with lower expression levels (176, 213)**.** While this would discourage the misfolding and aggregation of produced enzyme, it would also significantly reduce yield, which could hinder future crystallization attempts (176, 213). Another technique that could be beneficial is inclusion body solubilization and refolding (214, 215). This technique uses denaturants to solubilize the insoluble protein and dilution in a refolding buffer for refolding (214). If post-translational modifications were critical for the expression of these enzymes, heterologous expression systems such as yeast and baculovirus infected insect cells could allow proper expression (212). These systems are used less frequently than *E. coli* due to their more complex cloning and production of less total protein over a longer period of time (190). However, they are more likely to produce soluble insecticide resistance CBEs (212).

Chapter 3. The First Structural Characterization of an Insecticide Sequestering Carboxylesterase, Cqestβ2[1], from *Culex quinquefasciatus*

3.1. Journal article overview

As explained in the previous chapter, insecticide resistance is an ever-growing issue that endangers both our health and agricultural productivity (15, 97–99). The increased occurrence of resistance has encouraged extensive insecticide design, however, resistance remains an issue (103–106). Thus, new strategies for insecticide application and targets for insecticides need to be discovered to better combat resistance. While the number of insect species with resistance has been progressively growing, the number of resistance mechanisms remain relatively small (25, 126–128). By better understanding these mechanisms we will be able to determine new targets and formulate new strategies to combat insecticide resistance that may be broadly applicable.

One of the limiting factors in our understanding of insecticide resistance mechanisms has been the lack of molecular structures of the enzymes involved. Even in the most common insecticide resistance mechanism, CBE-mediated metabolic resistance, which relies on CBEs to either sequester or hydrolyze the insecticide before it reaches its target, only one enzyme's structure has been determined (5). This enzyme, LcαE7, acts by catalytically detoxifying insecticides and is thus an example of a qualitative resistance mechanism (5, 63, 161–163). Since its discovery this mechanism has been found in very few species, predominantly from the higher Diptera (161, 166, 167, 216). While the quantitative mechanism, insecticide sequestration, is more common and widespread, there was no enzyme structure determined until the work described in this chapter.

The southern house mosquito, *C. quinquefasciatus*¸ is not only an important vector of a range of filarial diseases including Japanese encephalitis, West Nile virus and Lymphatic filariasis but also utilizes CBE-mediated insecticide sequestration to provide resistance to a wide range of insecticides (4, 217, 218). In this chapter, we describe the expression and utilization of lysine methylation to crystallize and solve the structure of Cqestβ2[1], one of the most common insecticide resistance CBEs in *C. quinquefasciatus* (101, 219)**.**

To confirm that Cqestβ2[1] functions as an insecticide sequestration CBE with OPs and carbamates we conducted a range of inhibition assays between Cqestβ2[1] and relevant insecticides. Stopped-flow kinetic analysis demonstrated that the reaction with OPs proceeds via a rapid binding event with high affinity that results in an essentially irreversible, covalent intermediate. This agrees with previous studies and explains its action in insecticide sequestration (33). While its interaction was strong with OP insecticides, it demonstrated a weaker interaction with the carbamate tested suggesting a minor role in carbamate resistance.

To better understand the relationship between Cqestβ2[1], insecticide resistance CBEs and other insect CBEs we utilized a novel technique called sequence similarity networks (SSNs) (220, 221). This utilizes all-by-all BLAST rather than multiple sequence alignment allowing much larger sets of sequences to be analyzed than phylogenetic trees (thousands opposed to hundreds) (222). While this technique can generate groups of related sequences, it does not present any information on the evolutionary history of enzymes (222). Due to the increased sequence coverage, the risk of stochastic error present with phylogenies is reduced (223). The SSNs demonstrated that, irrespective of insect species, CBEs associated with insecticide resistance share a level of similarity that sets them apart from other insect CBEs. The SSNs also had the added benefit of revealing greater complexity to the insect CBE family suggesting further phylogenetic work may be required to improve classification.

To analyze the similarities and differences suggested by the SSN in more detail, we compared the structure of Cqestβ2[1] with other insect CBE structures that had been solved. This demonstrated key similarities between Cqestβ2[1] and the target of OP and carbamate insecticides, AChE, consistent with its function as an insecticide sequestration CBE (224). A comparison between Cqestβ2[1] and LcαE7 revealed both key similarities that enabled both to act as insecticide resistance CBEs and key differences that lead to each adopting different mechanisms (quantitative vs qualitative change, respectively) (5, 63). The evidence so far suggests that CBEs recruited through evolution for a role in insecticide resistance possess larger and less specialized binding pockets than hormone, odorant or neurotransmitter CBEs.

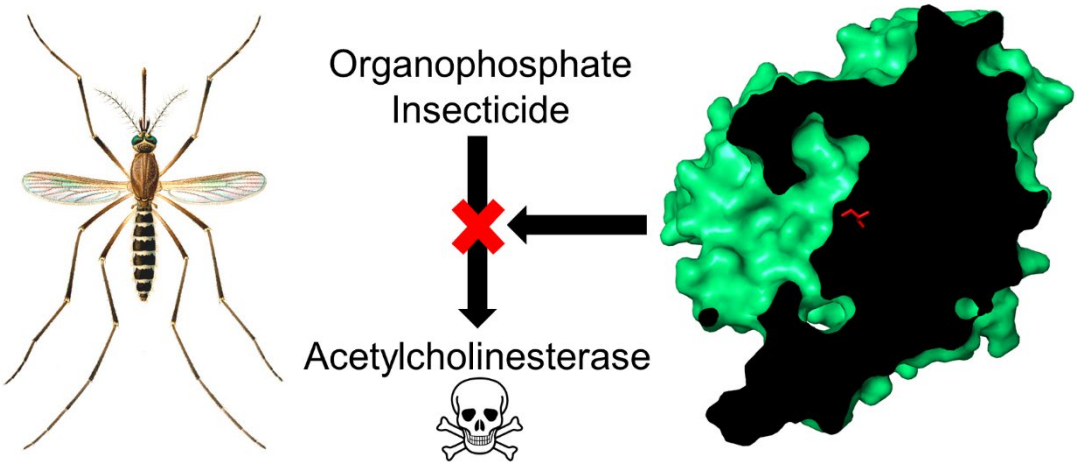
So far, 11 naturally occurring isoforms of Cqestβ2[1] have been identified (218). While many studies have implicated each in resistance, there has been no direct comparison of each isoforms reactivity with insecticides, stability, and/or optimization for expression in *Culex* mosquitoes (218). In this chapter we first compared the sequence diversity amongst the isoforms and found that Cqestβ1, possessed the most significant deviations from Cqestβ2[1]. Biochemical and stability studies demonstrated that the sequence diversity is mostly neutral in relation to OP binding and that Cqestβ2[1] possessed greater thermostability. This increased stability may be related to the wider distribution of Cqestβ2[1] over other isoforms (101, 219). The similarity in OP binding suggests future targeted inhibitor design against Cqestβ2[1] is likely to be broadly applicable.

3.2 Statement of contribution

Structure of an Insecticide Sequestering Carboxylesterase from the Disease Vector *Culex quinquefasciatus*: What Makes an Enzyme a Good Insecticide Sponge?

Davis H. Hopkins, Nicholas J. Fraser, Peter D. Mabbitt, Paul D. Carr, John G. Oakeshott and Colin J. Jackson

Biochemistry, 2017, 56, 5512 - 5525



This paper has been peer-reviewed and published as an original research article. All protein cloning, expression and purification, crystal screening, enzyme assays, construction of sequence similarity networks, molecular comparisons, CD spectroscopy, analysis of results, writing of the paper and discussion are my own work. Paul Carr collected diffraction data and, along with Colin Jackson, assisted in solving the protein crystal structure. Nicholas Fraser and Peter Mabbitt consulted on the appropriate methods for the experiments. Colin Jackson and John Oakeshott supervised the work, in addition to contributing to writing and editing the manuscript.

Relevant sections of the supporting information for this paper are presented following the manuscript.

Biochemistry

Article

Cite This: *Biochemistry* 2017, 56, 5512-5525

pubs.acs.org/biochemistry

Structure of an Insecticide Sequestering Carboxylesterase from the Disease Vector *Culex quinquefasciatus:* What Makes an Enzyme a Good Insecticide Sponge?

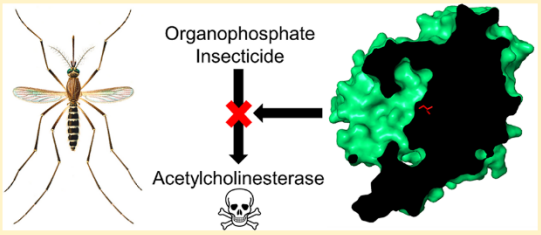
Davis H. Hopkins,[†] Nicholas J. Fraser,[†] Peter D. Mabbitt,[†] Paul D. Carr,[†] John G. Oakeshott,[‡] and Colin J. Jackson*[,†]

[†]Research School of Chemistry, Australian National University, Canberra, Australian Capital Territory 0200, Australia

[‡]CSIRO, GPO Box 1700, Canberra, Australian Capital Territory 2601, Australia

Supporting Information

ABSTRACT: Carboxylesterase (CBE)-mediated metabolic resistance to organophosphate and carbamate insecticides is a major problem for the control of insect disease vectors, such as the mosquito. The most common mechanism involves overexpression of CBEs that bind to the insecticide with high affinity, thereby sequestering them before they can interact with their target. However, the absence of any structure for an insecticide-sequestering CBE limits our understanding of the molecular basis for this process. We present the first structure of a CBE involved in sequestration, Cqestβ2[1], from the mosquito disease vector *Culex quinquefasciatus*. Lysine methylation was used to obtain the crystal structure of Cqestβ2[1], which adopts a canonical α/β-hydrolase fold that has high similarity to the target of organophosphate and carbamate insecticides, acetylcholinesterase. Sequence similarity networks of the insect carboxyl/cholinesterase family demonstrate that CBEs associated with metabolic insecticide resistance across many species share a level of similarity that distinguishes them from a variety of other classes. This is further emphasized by the structural similarities and differences in the binding pocket and active site residues of Cqestβ2[1] and other insect carboxyl/cholinesterases. Stopped-flow and steady-state inhibition studies support a major role for Cqestβ2[1] in organophosphate resistance and a minor role in carbamate resistance. Comparison with another isoform associated with insecticide resistance, Cqestβ1, showed both enzymes have similar affinity to insecticides, despite 16 amino acid differences between the two proteins. This provides a molecular understanding of pesticide sequestration by insect CBEs and could facilitate the design of CBE-specific inhibitors to circumvent this resistance mechanism in the future.



Chemical insecticides have been the mainstay of insect pest control in agriculture and public health for over 50 years.[1] This has resulted in selection for pest genotypes that are resistant to the insecticides, compromising their effectiveness.[2] Organophosphates (OPs) have been one of the most widely used classes of insecticides, and OP resistance has been identified in many agricultural insect pests and disease vectors, including aphids (e.g., *Aphis gossypii, Myzus persicae*),[3,4] blowflies (e.g., *Lucilia cuprina*),[5] and mosquitoes (e.g., *Aedes aegypti, Culex quinquefasciatus*).[6,7] Multiple insecticide resistances, including OP resistance, have been detected in *C. quinquefasciatus* (the southern house mosquito), which is an important vector of several diseases, including Japanese encephalitis, lymphatic filariasis, and West Nile virus, in many parts of the world.[8,9]

A major OP and carbamate resistance mechanism in *C. quinquefasciatus* and related mosquito species involves overexpression of metabolic enzymes such as carboxylesterases (CBEs), which are able to react with the insecticides. CBEs normally catalyze the hydrolysis of carboxylesters and thus share features in their active sites that promote this reaction. When a carboxylester reacts with a CBE its alcohol and acyl groups are stabilized in specific subpockets that orient the ester ideally for reaction with the catalytic serine. This reaction results in the formation of a tetrahedral intermediate that is stabilized by an oxyanion hole through hydrogen bonding with the amino groups of a number of glycine residues. The alcohol group is released from the substrate forming an acyl-enzyme complex. Water then rapidly interacts with this complex forming a new tetrahedral intermediate that collapses, releasing the carboxylic acid group and regenerating the catalytic serine for further reaction.[10] In the case of OP and carbamate insecticides, the catalytic serine reacts with the insecticide in a similar manner to the formation of the acyl-enzyme complex with the release of an alcohol group. However, the reacted enzyme is highly stable, and regeneration with water is often slow and in some cases does not occur, resulting in irreversible inhibition (Figure S1).[11] This covalent conjugation results in the insecticide being stoichiometrically sequestered before it

Received: August 11, 2017
Revised: September 7, 2017
Published: September 20, 2017

DOI: 10.1021/acs.biochem.7b00774
Biochemistry 2017, 56, 5512−5525

reaches its target, acetylcholinesterase (AChE), a vital enzyme in the insect central nervous system.[12−14] This sequestration mechanism has been found in many insects, but has been most thoroughly studied in *M. persicae* and various culicine mosquitoes, including *C. quinquefasciatus*.[4,7] All *Culex* species known to have metabolic resistance to OPs share the same CBE-mediated sequestration mechanism, and the CBEs involved are highly conserved in terms of sequence identity.[15]

Two types of CBE are overexpressed in OP-resistant *C. quinquefasciatus*, Estα and Estβ, which are encoded by two tightly linked loci, *Est*-3 and *Est*-2, respectively.[16] These two types share around 50% amino acid sequence identity but exhibit limited immunological cross-reactivity.[17] As the various alleles at the two loci often occur in complete linkage disequilibrium they are referred to as alleles of a single super locus (named *Ester*).[18] To date, 13 different *Ester* alleles from *Culex* species have been associated with overexpression of either one or both of the encoded CBEs.[15] It has been found that these CBEs are commonly expressed in the digestive system but that expression patterns of the same alleles can vary greatly among *Culex* strains. Some strains have been found to have expression localized to the nervous system.[19] The allele *Ester2*, encoding estα2[1] and estβ2[1], is the most widespread and most commonly associated with OP resistance.[16,18] Overexpression is caused by a combination of both upregulation and gene amplification.[20] One study comparing the gene amplification of field populations of *Culex pipiens* to a standard laboratory OP susceptible strain found that the level of estα2[1] and estβ2[1] gene amplification ranged from 3- to 22-fold in the field populations.[21] In *C. quinquefasciatus* the equivalent CBEs have both been isolated and have also been shown to react with OPs, consistent with a role in resistance.[22] *C. quinquefasciatus* estβ2[1] (Cqestβ2[1]) has been associated with resistance to a wide variety of OPs (Naled, temephos, malaoxon, etc.).[9,14,23] It also has been implicated in resistance to other insecticide classes including carbamates (propoxur, Fenobucarb), which also target AChE, and synthetic pyrethroids (SPs) (resmethrin, permethrin, tetramethrin), which target voltage gated sodium channels.[8,14,24] Carbamates contain an amide bond and are generally not hydrolyzed by CBEs, while SPs contain an ester bond and can be hydrolyzed by some.[25]

Despite extensive analysis of the biology, genetics, and role of Cqestβ2[1] in insecticide resistance, the absence of a molecular structure has prevented us from understanding its interactions at a molecular level. In this work, we describe the first heterologous expression of Cqestβ2[1] and the utilization of lysine methylation to promote crystallization, allowing the structure to be determined via X-ray crystallography. This enabled a comprehensive structural comparison between the active site of Cqestβ2[1] and closely related CBEs including the insecticide target, AChE. Sequence similarity networks (SSNs) are used to analyze the relationships between 3162 different insect CBEs, to aid in their classification and determine the relationship of Cqestβ2[1] to other insect CBEs. Finally, activity and inhibition profiling with substrates and insecticides is used to define their specificity and demonstrate the similarity of Cqestβ2[1] to its isoforms through comparison with one of its most divergent isoforms, Cqestβ1.

■ MATERIALS AND METHODS

Cloning, Protein Expression and Purification of Cqestβ2[1] and Cqestβ1. The Cqestβ2[1] gene (Genbank accession number CAA83643)[26] and Cqestβ1 gene (Genbank accession number AAA28289.1)[27] of *C. quinquefasciatus* were synthesized with an N-terminal His-tag and TEV cleavage site and optimized for expression in *Escherichia coli*. The genes were then cloned into the vector pETMCS III through the use of Gibson assembly (New England Biolabs (NEB)).[28] Expression was conducted in *E. coli* BL21(DE3) (Invitrogen) cells grown at 25 °C for 20 h in autoinduction media supplemented with 100 μg/mL ampicillin. Cells were pelleted at 5000*g* and lysed by sonication in lysis buffer (50 mM HEPES (pH 8.0) and 500 mM NaCl). Lysate was filtered and passed through a 5 mL HisTrap FF column (GE Healthcare) before being eluted by lysis buffer supplemented with 250 mM Imidazole. The His-tag was cleaved through a TEV reaction by exchanging protein into TEV reaction buffer (50 mM TRIS-HCl (pH 8.0), 300 mM NaCl, 0.5 mM EDTA, 1 mM DTT) containing TEV protease at a 1:20 ratio with calculated protein concentration and reacting overnight at 18 °C. The protein solution was then passed through the HisTrap FF column, separating cleaved from His-tagged protein, before being purified by size exclusion chromatography (SEC) using a HiLoad 16/600 Superdex 75 PG (GE Healthcare) in 20 mM HEPES (pH 7.5) and 150 mM NaCl.[29] Protein concentrations were estimated by measuring absorbance at 280 nm using an extinction coefficient of 75 915 M^{-1} cm^{-1} calculated using the Protparam server.[30] Protein purity was confirmed through SDS-PAGE analysis.

Crystallization and Structure Determination of Cqestβ2[1]. Cqestβ2[1] was modified through lysine methylation using the protocol of Walter et al. (2006)[31] to improve crystallization. In brief, solutions of concentrated dimethylamine-borane complex and formaldehyde were sequentially added to the protein solution before incubation overnight at 4 °C. Precipitation was removed through filtration, and protein was purified through SEC. A single crystal was obtained using the sitting drop diffusion method using reservoir solutions of 1.6 M ammonium sulfate and 0.1 M HEPES (pH 7.0) with a Cqestβ2[1] concentration of 12 mg/mL. The crystal was then flash-cooled in liquid nitrogen using the cryoprotectant 2 M sodium malonate. Diffraction data were collected on the MX2 beamline at the Australian Synchrotron, Victoria, Australia, with a wavelength of 0.95370 Å.[32] The data were indexed, integrated, and scaled using MOSFLM.[33] Data collection statistics are shown in Table 1. Phases were obtained through molecular replacement using the structure of LcαE7,[13] and the program MOLREP.[34] Model rebuilding was done through iterative cycles of automated rebuilding using ARP/wARP[35] and manual rebuilding using COOT.[36] Refinement was done using REFMAC[37] accessed through the CCP4 suite of programs[38] and PHENIX.[39] There are a total of eight Ramachandran outliers in the structure that occur in areas of poor electron density.

Enzyme Assays. All assays were performed at room temperature.

Nitrophenyl Substrates. Assays were performed in triplicate using 200 μL reactions with a range of substrate concentrations from 0 to 1 mM in 20 mM HEPES (pH 7.5), 150 mM NaCl, and 2% methanol. Enzyme concentration varied between 5 nM and 100 nM. Product formation was measured at 405 nm using a microplate spectrophotometer. Concentration of product was determined by using a molar extinction coefficient (ε = 18 400 M^{-1} cm^{-1}). Kinetic parameters were determined by fitting the initial velocity data to the Michaelis−Menten equation using nonlinear regression. Because of its limited solubility, an

DOI: 10.1021/acs.biochem.7b00774
Biochemistry 2017, 56, 5512−5525

Table 1. X-ray Data Collection and Refinement Statistics

data collection	Cqestβ2[1]
space group	$P3_221$
unit cell parameters	
a, b, c (Å)	125.07, 125.07, 168.81
α, β, γ (deg)	90, 90, 120
wavelength (Å)	0.95370
resolution range (Å)[a]	45.58−2.50 (2.58−2.50)
no. of unique reflections	53392 (4561)
redundancy	10.7 (10.9)
completeness (%)	100 (100.0)
$R_{merge(I)}$[a]	0.161(1.226)
R_{pim}	0.051 (0.387)
mean $<I/\sigma(I)>$	11.7 (2.2)
$CC_{1/2}$	0.996 (0.578)
Refinement	
no. reflections (total)	53336
resolution range (Å)	45.581−2.50
R_{work}/R_{free} (%)	19.47/26.98
no. of macromolecules	2
no. of total atoms	
protein	16831
ligands	82
water	260
B factor (Å^2)	
protein	55.7
ligand	64.9
water	36.9
Ramachandran plot (%)	
favored regions	90
allowed regions	9
outlier regions	1
RMSD	
bond lengths (Å)	0.016
bond angles (o)	1.646
Protein Data Bank ID code	5W1U

[a]Values in parentheses are for the highest-resolution shell.

apparent k_{cat}/K_M was estimated for 4-nitrophenyl octanoate (eq 1).[13]

$$k_{cat}/K_M^{app} = v_o/([E]/[S]) \quad (1)$$

Diethyl 4-Methylumbelliferyl Phosphate. Stopped flow kinetics were determined through the use of a rapid-mixing stopped-flow unit coupled to a UV/vis spectrophotometer. In each assay DEUP concentration was varied between 1.5 and 48 μM, and enzyme concentration was 0.75 μM with at least three replicates per concentration. Assays were performed in 20 mM HEPES (pH 7.5), 150 mM NaCl, and 2% methanol. Hydrolysis of DEUP was monitored fluorometrically (excitation 330 nm, emission 450 nm). Kinetic parameters were determined as by Mabbitt et al. (2016) based on the proposed scheme for hydrolysis of DEUP (eq 2).[40]

$$E + S \underset{k_2}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_3} EPQ \xrightarrow{k_5} E + P + Q \quad (2)$$

Cypermethrin Analogues. The synthesis and structure of the analogues of the cypermethrin diastereomers have been previously described.[41] Assays were performed in triplicate using a single concentration of each analogue (10 μM) due to their limited solubility in water. Enzyme concentration was varied from 0.02 μM to 0.10 μM, and all assays were performed in 20 mM HEPES (pH 7.5), 150 mM NaCl, and 0.5% DMSO. The analogues used are hydrolyzed into a cyanohydrin product that spontaneously forms the fluorescent 6-methoxynaphthaldehyde. Thus, the reaction was followed by monitoring 6-methoxynaphthaldehyde formation (excitation 315 nm, emission 460 nm).

Inhibition Kinetic Constants. Two methods were used to determine inhibition kinetic constants with a range of inhibitors: For the insecticides temephos and propoxur, the constants were determined by first incubating purified enzyme with a range of inhibitor concentrations. Then at various times aliquots were taken, and residual activity was determined using the substrate, 4-nitrophenyl acetate. The reaction rate constants were determined using the methods of Aldridge and Reiner.[42] For paraoxon-ethyl, inhibition assays were performed in the presence of substrate, and reaction rate constants were determined using the method of Hart and O'Brien.[43] Briefly, the purified enzyme was added to reactions with a set concentration of 4-nitrophenyl acetate or butyrate and varying concentrations of inhibitor. The rate of inhibition was monitored through the loss of activity with 4-nitrophenyl substrate as stated above. All measurements were performed in triplicate. Both methods are based on the following inhibition scheme (eq 3).[43]

$$E + PX \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} EPX \xrightarrow{k_2} EP + X \quad (3)$$

Computational Tools for Protein Visualization and Sequence Similarity Networks. Pymol was used for all structure comparison and protein images[44] except for the topology diagram produced using TOPDRAW.[45] The automated PyMOL representation of electrostatics, protein contact potential, was used for qualitative comparison of electrostatic properties. Multiple sequence alignment was performed using MEGA6 and visualized using BoxShade. The sequence similarity networks were generated using the CBE family (Pfam PF00135, 3162 sequences including sequences from the following insect orders: Coleoptera, Diptera, Hemiptera, Hymenoptera, Isoptera, Lepidoptera, Orthoptera, and Phthiraptera), including enzymes of particular interest (Supporting Information List S1) and the Enzyme Function Initiative Enzyme Similarity Tool (EFI-EST, http://efi.igb.illinois.edu/efi-est/).[46] Sequences from the orders Blattodea and Decapoda were not included in the SSNs as they each formed less than three nodes and were thus uninformative. An expectation value of 10^{-50} was used to restrict the all-by-all BLAST. The network was visualized through the use of Cytoscape v 3.2.1.[47] Each node represents sequences with at least 90% sequence identity, and edges represent the similarity between these nodes. Similarity is measured using a database-independent metric developed by EFI-EST called alignment scores, which is based on the bit-score used in BLAST searches, where the greater the score the more similarity between sequences.[46] Two networks were generated one with an alignment score of 74 and the other of 93.

CD Spectroscopy and Thermal Denaturation Curves. CD spectra were recorded using an Applied Photophysics Chirascan spectropolarimeter connected to a temperature control unit. Proteins were diluted into phosphate storage buffer (20 mM NaH_2PO_4, 50 mM NaCl, pH 8.0) at a concentration of 0.5 mg/mL. CD spectra were initially measured at 1 nm intervals between 200 and 260 nm at 20 °C. Bandwidth was set to 1 nm and the scan length to 0.5 s per

DOI: 10.1021/acs.biochem.7b00774
Biochemistry 2017, 56, 5512−5525

point. Thermal denaturation of proteins was observed by recording CD at 208 nm at 0.5 °C intervals as the temperature was increased from 20 to 90 °C at a rate of 1 °C·min^{-1} with a scan length of 3 s per point. Experiments were performed in triplicate. Spectra were taken after denaturation at 90 and 20 °C to confirm denaturation and lack of refolding.

RESULTS AND DISCUSSION

Structure of Cqestβ2[1]. Expression of the Cqestβ2[1] gene in the BL21 (DE3) strain of *E. coli* was found to produce soluble and active Cqestβ2[1] protein. Initial crystallization trials at varying protein concentrations did not yield diffraction-quality crystals, but lysine methylation, a technique that improves protein−protein interactions and crystallization, resulted in a single crystal in 1.6 M ammonium sulfate and 0.1 M HEPES (pH 7.0) that diffracted to 2.5 Å resolution. The structure of Cqestβ2[1] was then determined by molecular replacement using the structure of the αE7 CBE from *L. cuprina*[13] (LcαE7, PDB ID: 4FG5; 37% amino acid identity) as a template (Table 1). The crystal structure contains two monomers per unit cell. Residues 5−540 of the A monomer and 4−540 of the B monomer are well resolved in clear electron density. The PISA server does not detect any plausible dimer interfaces.[48] This was confirmed through SEC, which resulted in a single peak at an elution volume consistent with a monomer.[49]

As shown in the topology map of Cqestβ2[1], the structure adopts the α/β-hydrolase superfamily fold (Figure 1A). This consists of an eight-stranded β-sheet surrounded by six α-helices.[50] The canonical β-sheet is extended by the addition of one strand at the N-terminus and two antiparallel strands at the C-terminus, while the canonical α-helices D and F are both separated into two shorter helices. The location of the catalytic triad is conserved with most members of the α/β-hydrolase superfamily: the nucleophilic serine, located at the end of a sharp turn after β-strand 5, the glutamic acid, occurring after β-strand 7, and the histidine after β-strand 8. The substrate-binding pocket is defined predominantly by subdomains 1 and 2, which extend above the core fold and form a cleft that surrounds the catalytic residues (Figure 1A,B). Subdomain 1 contains a short antiparallel β-sheet consisting of two strands and a cluster of three α-helices that occurs between β-strand 1 and α-helix D. This subdomain is shifted away from the catalytic residues with only the residues F281, W224, L120, Y121, and G108-11 in close proximity to it (Figure 3A). Subdomain 2 contains a cluster of five α-helices that occur between β-strand 7 and α-helix F. This subdomain contributes the majority of residues that define the binding pocket including F453, N452, M431, L446, Y428, Q330, H442, L327, L328, F394, and V393 (Figure 3A).

Relationship of Cqestβ2[1] to Other Insecticide Resistance CBEs. Insect carboxylesterases, including cholinesterases (the CBE family), have been classified into 14 major clades (A−N) based on comprehensive phylogenies—principally using the genomes of seven insect species from the orders Diptera, Hymenoptera, and Coleoptera.[1,51,52] The topology of the phylogeny was retained when the genomes of additional hymenopterans were added.[53] The 14 clades were grouped into three broad classes based on a more general description of function: dietary/detoxification, hormone/semiochemical processing, and neuro/developmental functions.[52,53] Phylogenies using several additional genomes, including Lepidoptera and Hemiptera, broadly support these classes, although they also reveal some additional complexity, particularly in the dietary/detoxification and hormone/semiochemical processing classes.[54−60]

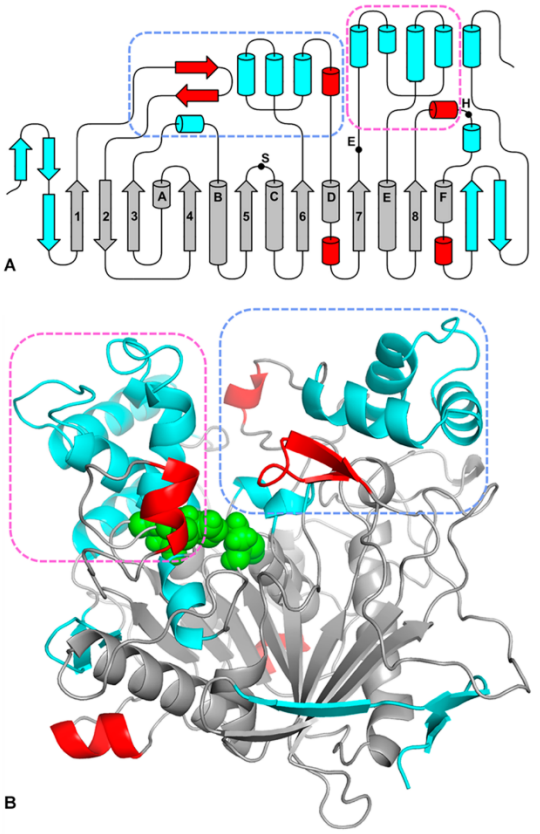


Figure 1. (A) A topology diagram of Cqestβ2[1] demonstrating the core, canonical α/β-hydrolase fold (gray) and secondary structure absent (red) and shared (blue) with *Anopheles gambiae* AChE-1. The two binding domains, subdomain 1 (boxed in blue) and subdomain 2 (boxed in pink), are also shown. The active site residues are represented as dots and labeled with their amino acid using the single letter code. (B) A cartoon diagram of Cqestβ2[1] facing the binding pocket entrance and demonstrating the canonical α/β-hydrolase fold with secondary structure labeled as above. The catalytic residues are represented as spheres (green).

Sequence similarity networks (SSNs) are a comparatively new method for visualizing and inferring functional relationships in large protein families that can be useful in classification.[61] So far, SSNs have not been used to analyze the insect CBE family. Unlike phylogenetic trees, which use multiple sequence alignment, SSNs compare sequences through an all-by-all BLAST. This allows the incorporation of a much larger group of sequences than phylogenetic trees (thousands opposed to hundreds), which reduces the risk of stochastic error.[62] Trends in the set of sequences can be easily observed by including relevant information for each individual protein that can be mapped onto networks (functional annotation, species, etc.) and visualized.[47,62] Unlike phylogenetic trees, SSNs do not infer phylogeny and only demonstrate groups of similar sequences that may suggest similar functions.[62,63] We used SSNs to better understand the

DOI: 10.1021/acs.biochem.7b00774
Biochemistry 2017, 56, 5512−5525

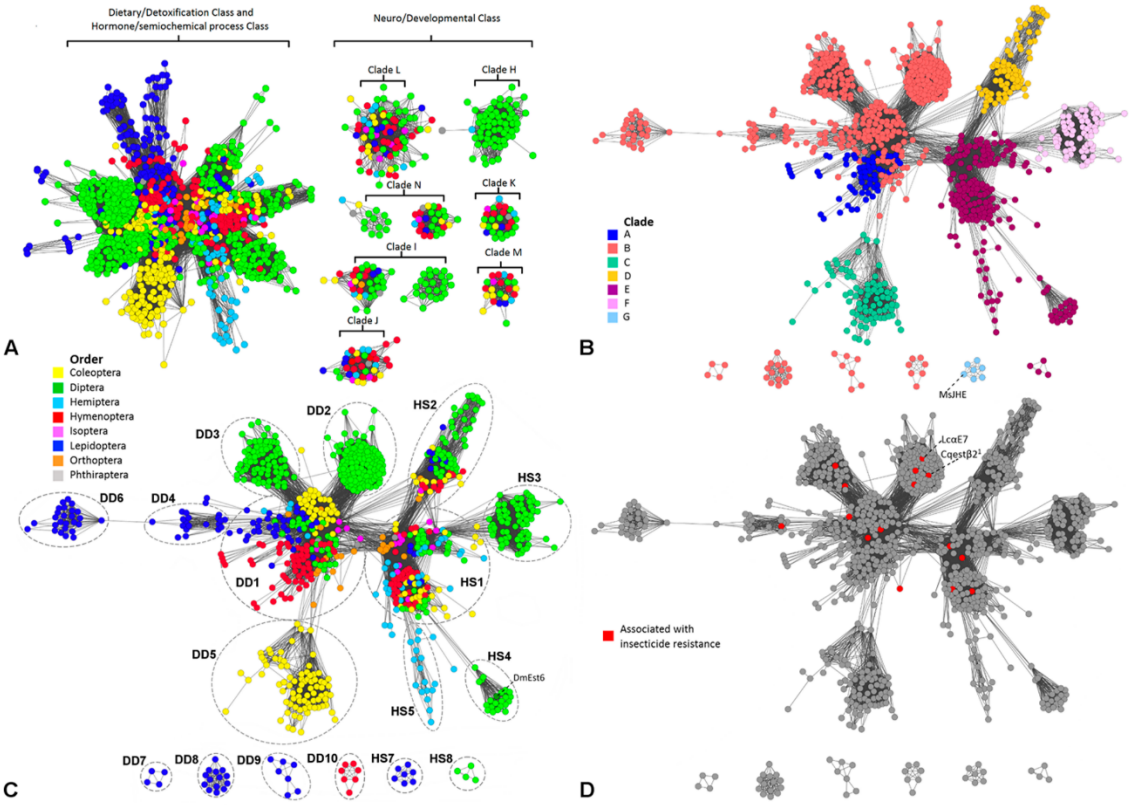


Figure 2. SSNs of the insect CBE family where each node represents proteins sharing 90% sequence similarity and edges represent the alignment score between proteins. (A) A SSN using an alignment score cut off of 74 with nodes colored based on insect order and classes and clades labeled. All subsequent SSNs use an alignment score cut off of 93 and exclude nodes from the neuro/developmental class: (B) A SSN with nodes colored based on a previous classification of clades.[52,53] (C) A SSN with nodes colored based on insect order and groups circled and labeled based on class (DD for detox/detoxification and HS for hormone/semiochemical processing). (D) A SSN highlighting nodes that contain insecticide resistance CBEs (red).

sequence diversity and classification of the insect CBE family and determine the relationship of Cqestβ2[1] to the other insecticide resistance CBEs.

The first SSN only includes edges with an alignment score of greater than 74 (Figure 2A). This separates the CBEs that were previously defined in the neuro/developmental class from those in both the dietary/detoxification and hormone/semiochemical processing classes. Each clade within the neuro/developmental class is clearly defined, supporting the previously suggested classifications (clade H, glutactins, clade I, unknown functions, clade J, AChEs, clade K, unknown function, clade L, neuroligins, clade M, gliotactins, clade N, neurotactins). Interestingly, clades N and I are each divided into two distinct groups: one consisting of predominantly dipteran CBEs and the other consisting of CBEs from all of the other sampled insect orders.

Higher stringency SSNs were then created to resolve the dietary/detoxification and hormone/semiochemical processing classes (Figure 2B). Here, nodes are colored based on the previous classification of the clades within the two classes.[52] In the hormone/semiochemical processing class, clade D (integument esterases), clade F (dipteran-type juvenile hormone esterases (JHEs)), and clade G (lepidopteran-type JHEs), all cluster into clear groups supporting the previous classification, groups HS2, HS3, and HS7, respectively.[52,53] Clade G includes *Manduca sexta* JHE (MsJHE), one of the few structurally characterized CBEs.[64] The previous classification of the rest of the clades in this class is less consistent with the remaining groups. Clade E (secreted β-esterases) loosely forms four different groups, instead of one: HS1, a general group including all insect orders, HS4, a clearly differentiated group of solely dipteran CBEs, which includes the odorant degrading enzyme Est6 from *D. melanogaster*,[65] HS5, a closely related group of hemipteran CBEs, and HS8, a small and less related group of dipteran CBEs (Figure 2B,C). Two of the earlier phylogenetic analyses focusing on either dipteran or hemipteran CBEs also differed from the latest classification, both suggesting that clade E existed as more than one group.[56,58] This difference is probably due to the sequences sampled, with the SSN including sequences from more divergent groups than in the previous classifications, and suggests greater functional diversity.

The phylogenies in Oakeshott et al. (2010) and Sadd et al. (2015) resolved the dietary/detoxification class into clade A, containing hymenopteran xenobiotic metabolizing enzymes, and clades B and C, which are more widely distributed across insect orders.[52,53] The large clade B, α-esterase type enzymes,

DOI: 10.1021/acs.biochem.7b00774
Biochemistry 2017, 56, 5512−5525

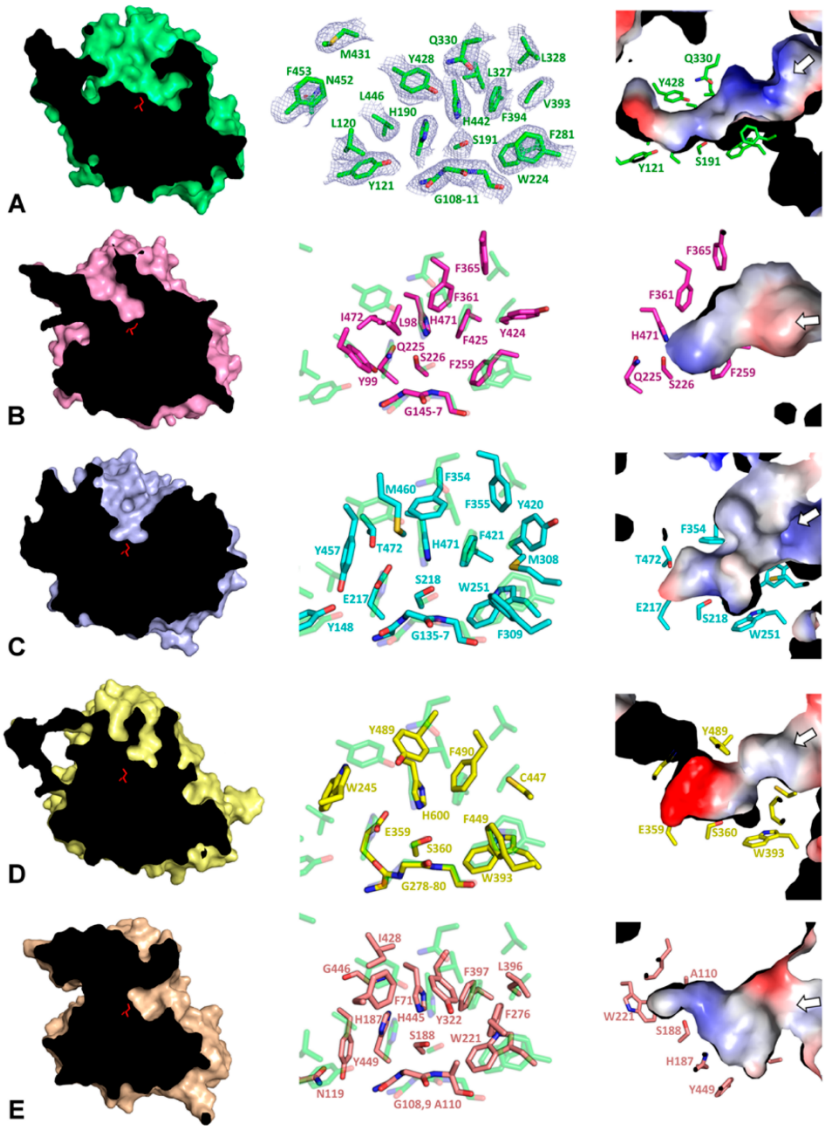


Figure 3. A comparison of the binding pocket orientation, active site residues and pocket electrostatic properties of (A) Cqestβ2[1] (green), (B) MsJHE (pink), (C) LcαE7 (blue), (D) AgAChE-1 (yellow), and (E) DmEst6 (tan). The binding pockets are first shown as surface representations that have been split vertically by a plane parallel to the page. The active site residues of Cqestβ2[1] are shown with the $2|F_O - F_C|$ map represented as a gray mesh contoured at 0.9σ (A) and as a transparent alignment with all other active site residues (B−E). The electrostatic properties of the binding pockets are shown in similar orientations to the active sites with the exception of DmEst6 which is rotated and viewed from the other side emphasizing its alternate entrance. Charge is represented as red (negative), white (no charge), and blue (positive) with saturation indicating strength. Arrows indicate the direction of substrate entry.

contains the Cqestα2[1] and Cqestβ2[1] CBEs, as well as LcαE7, while clade C, titled unknown functions, consists of CBEs yet to be functionally characterized. In contrast, the SSN presented here, which is the most comprehensive bioinformatic analysis in terms of species coverage, shows 10 clearly differentiated groups for the dietary/detoxification class. Of these, there is only one general group, DD1, which contains representatives of all insect orders sampled (Figure 2B,C). Four order-specific groups are directly linked to this group, DD2−5. DD2 contains only dipteran esterases including LcαE7 and Cqestβ2[1], both involved in lipid metabolism. DD3, dipteran esterases, and DD5, coleopteran esterases, contain no esterases with specific annotation, while DD4, lepidopteran esterases, is made up predominantly of antennal esterases. DD7−9 are all small groups of lepidopteran esterases without functional annotation. Each of these groups are disconnected from any other, suggesting distinct roles. DD10 is a group of only hymenopteran esterases without functional annotation. This

DOI: 10.1021/acs.biochem.7b00774
Biochemistry 2017, 56, 5512−5525

division in the dietary/detoxification class suggests that the functions of these proteins are likely to be more specific and differentiated than previously thought. Clearly, more work in this field is required for functional annotation of this complex family in insects.

Figure 2D shows which groups contain CBEs associated with metabolic insecticide resistance (List S2). In the dietary/detoxification class, these CBEs occur either in DD1 or in four of the five groups that are directly linked to it (HS1 and DD2−4) with none found in the unlinked, order-specific groups (DD7−10) or the linked coleopteran and lepidopteran groups (DD5 and DD6, respectively). The dipteran insecticide resistance CBEs, Cqest$\beta2^1$ and LcαE7, are both found in DD2. Previous studies have identified CBEs that, upon mutation, are able to catalytically hydrolyze OPs rather than sequester them.[66,67] These CBEs are largely restricted to DD1 and 2, while the sequestering CBEs are more broadly distributed. Notably, there are a number of insecticide resistance CBEs in the general β-esterase group, HS1, within the hormone/semiochemical processing class.[52] As HS1 is linked to DD1 it shows a similarity in sequence that may suggest a similarity in structure and/or function that allows both to adopt roles in insecticide resistance and may distinguish HS1 from other groups in Clade E like the odorant degrading CBE group, HS4, containing DmEst6. It is also consistent with transcriptomic evidence from both *D. melanogaster* and the lepidopteran *Spodoptera littoralis*, in which several CBEs are expressed in both semiochemical processing and detoxification tissues (the antennae and the gut, respectively).[68,69]

The SSNs have shown that there is greater sequence and functional diversity in the insect CBE family than previously thought. The distribution of insecticide resistance CBEs demonstrates that while these CBEs can come from a range of functional groups, they share a level of similarity that distinguishes them from other CBEs and allows them to react with the current range of commercially used insecticides. This suggests that insecticides need to be more specific to their targets to circumvent this similarity and that the development of synergists with broad inhibition of these CBEs could be possible.

Comparison with Other Structurally Characterized Insect CBEs. To better understand how the structure of Cqest$\beta2^1$ allows it to confer metabolic resistance to insecticides and why some of the other groups of CBEs are not predisposed to being involved in metabolic insecticide resistance, we compared it to the structures of all other structurally determined insect CBEs: AgAChE-1, the native AChE-1 in *A. gambiae*, is the primary target of organophosphate and carbamate insecticides in *A. gambiae* and representative of other insect AChE-1s (PDB ID: 5X61);[70] LcαE7, from *L. cuprina*, is the only other insecticide resistance CBE to be structurally determined (PDB ID: 4FNM);[13] MsJHE, from *Manduca sexta*, is the only JHE to be structurally characterized and is critical in regulating insect development (PDB ID: 2FJ0);[64] DmEst6, from *Drosophila melanogaster*, is the first structure of an odorant-degrading CBE to be determined and plays a role in chemosensory systems (PDB ID: 5THM).[65] Comparison with the structure of AChE-2 from *D. melanogaster* (PDB ID: 1QO9) was also undertaken but found to have only minor differences to AgAChE-1 and was thus excluded from this discussion.[71]

LcαE7 has the closest structural similarity to Cqest$\beta2^1$ of any known structure, showing an RMSD of 1.41 Å, even though its amino acid sequence is relatively different (37% sequence identity).[72] The SSN in Figure 2D showed that both Cqest$\beta2^1$ and LcαE7 lie in a single group limited to CBEs from Diptera, DD2. One major difference is the presence of an extended N-terminal helix found in LcαE7 but not Cqest$\beta2^1$, which was suggested to be involved in membrane/lipid association and cellular localization of LcαE7.[13] DmEst6, AgAChE-1, and MsJHE also show high similarity with Cqest$\beta2^1$ in terms of the overall fold with Cα RMSDs of 1.94, 1.56, and 2.07 Å, respectively, while having sequence identities of 25%, 33% and 29%, respectively.[72] A comparison of Cqest$\beta2^1$ with the topology and structure of AgAChE-1 shows that the majority of secondary structural elements are shared (shown in light blue, Figure 1A). There are only five short secondary structural elements that differ: three in the subdomains of Cqest$\beta2^1$ and two in its core structure (shown in red, Figure 1).

The size and shape of each structure's substrate binding site vary in line with their respective substrates (Figure 3 and Table S1 for substrate structures). The CASTp server was used to predict binding pocket volumes for better comparison.[73] AgAChE-1 possesses a small, asymmetrical pocket with a long tunnel (predicted volume of 1736.2 Å^3) that is specialized for its substrate, acetylcholine. MsJHE has a long, narrow tunnel (predicted volume of 1308 Å^3) that suits its substrate, juvenile hormone (JH), an α,β unsaturated ester with a long hydrophobic acyl group. DmEst6 has a relatively small pocket (predicted volume 935 Å) that reacts with a range of short-mid chain food esters.[65] LcαE7 has a larger pocket (predicted volume of 2727.5 Å^3), which has been shown to be capable of hydrolyzing large substrates such as fatty acid methyl esters (FAMEs), consistent with a role in lipid metabolism as suggested by knockout studies.[13,74] Cqest$\beta2^1$ has the largest binding pocket of all of these CBEs (predicted volume of 4735.5 Å^3), consistent with evidence that it reacts with medium-chain-length mono- and diacylglycerols and is involved in lipid-fatty acid metabolism.[12]

The active site entrance is mainly defined by two subdomains in Cqest$\beta2^1$ and LcαE7, with one subdomain preceding α-helix D and the other following β-strand 7. For both Cqest$\beta2^1$ and LcαE7, one boundary of the pocket is defined by a large loop that includes an antiparallel β-sheet following β-strand 1. This provides stability to this region and in LcαE7 is shown to be conformationally coupled to motions within the binding site that facilitate substrate binding.[40] At this position in MsJHE, there is an extended loop, while in AgAChE-1 there is a small helix and large loop insertion. Both these structural elements restrict the binding pockets in these proteins. In DmEst6 this loop adopts quite a different position, resulting in a unique entrance not shared with the other esterases. In Cqest$\beta2^1$ and LcαE7, the other side of the pocket is expanded by the formation of a small α-helix that separates the two subdomains. In both AgAChE-1 and DmEst6, this is replaced by a loop without secondary structure that extends toward the active site. In AgAChE-1 this restricts its shape but in DmEst6 this completely blocks access from this angle. In MsJHE there is no equivalent helix to the one in Cqest$\beta2^1$, but the helix that precedes this extends far further into the binding pocket, completely blocking access from this angle. The absence of the equivalent helix in MsJHE also allows the subdomains to collapse further over the binding pocket forming a narrow tunnel that has a unique entrance not shared with the other enzymes. Cqest$\beta2^1$, MsJHE, DmEst6, and LcαE7 all have a helix and large loop insertion after β-strand 8 that is not found

DOI: 10.1021/acs.biochem.7b00774
Biochemistry 2017, 56, 5512−5525

in AChE. This provides a boundary to the binding pocket that is already present in AChE in the form of a small α-helix and loop insertion following β-strand 1. In DmEst6 this feature is shifted to further define its unique entrance. Even though the c-α RMSD of each of these proteins to Cqestβ2[1] is quite low, the small differences in secondary structure can lead to greatly different binding pockets.

Active Site Comparison and Implications for Activity. The active site residues of Cqestβ2[1], LcαE7, MsJHE, AgAChE-1, and DmEst6 are shown in Figure 3. All have the canonical features shared by many other α/β hydrolases including a catalytic triad and an oxyanion hole. However, all these enzymes have a Ser-His-Glu catalytic triad and an oxyanion hole formed by three glycine residues except for DmEst6, which has a Ser-His-Asp triad and oxyanion hole formed by two glycine residues. The latter configuration of the catalytic triad is more common in the α/β hydrolase superfamily overall.[75]

The acyl-binding pockets are all predominantly made up of bulky, hydrophobic residues resulting in the binding pockets having largely nonpolar or lightly charged properties. In AgAChE-1 this pocket is restricted by F490, F449, C447, and W393 creating a small pocket close to the catalytic residues (Figure 3D). In both LcαE7 and Cqestβ2[1] the backbone is shifted significantly expanding the acyl-binding pocket. In Cqestβ2[1] replacement of some of the bulky residues with smaller alternatives such as L327, L328, and V393 further expands this pocket relative to LcαE7 (Figure 3A,C). Both also have positively charged residues on the outer edges of the pocket. In the crystal structure of Cqestβ2[1], a malonate ion is bound in the acyl binding pocket with its bridging methylene associated with the residues F394 and L327. In MsJHE this pocket is far less crowded because it forms a major part of its unique entrance (Figure 3B). This differs greatly to DmEst6, which is the most crowded, with residues such as F276 and Y322 shifted even further into this pocket (Figure 3E).

The alcohol-binding pocket differs greatly between Cqestβ2[1] and the other enzymes. In AgAChE-1 the pocket contains a number of polar residues (E359, Y489, H600) that result in a highly, negatively charged pocket and its size is predominantly restricted by one residue, W245. Both the charge and cation π-interactions with W245 are critical in binding the positively charged choline group of acetylcholine.[76] In both MsJHE and LcαE7 this pocket is small, with some polar residues close to the catalytic serine and nonpolar residues along with a tyrosine residue restricting one side greatly. This results in both possessing largely nonpolar pockets with areas of weak positive charge near the catalytic serine. In Cqestβ2[1] there is no tryptophan, tyrosine, or equivalent residue restricting the size of this pocket, allowing it to expand into a much bigger cavity made up predominantly of hydrophobic residues. Near the catalytic serine the pocket has a weak positive charge while further in the pocket develops a weak negative charge. This increased size in turn expands the binding pocket entrance, presumably increasing substrate diffusion into the pocket and potentially affecting the size of substrates with which it best interacts. A HEPES ion is bound into this pocket in the crystal structure with its hydroxyl group associated with the catalytic serine and its carbon backbone extending to L120 at the back of the pocket. This suggests a possible binding mode for esters with long, hydrophobic alcohol groups. DmEst6 also lacks a bulky residue restricting this side of the pocket. Instead this side is open to the solution, forming its unique entrance, which is lined with mostly hydrophobic residues. The entrance has a weak positive charge near the catalytic serine and a weak negative charge closer to the entrance.

To test how the binding pocket of Cqestβ2[1] interacts with substrates of varying size, three nitrophenyl ester substrates with different acyl groups were assayed (Table 2 and Table S1 for substrate structures). Cqestβ2[1] showed relatively high specificity constants (k_{cat}/K_M) for all substrates. The highest specificity was found with 4-nitrophenyl butyrate. This is surprising given that the large acyl binding pocket would suggest better interactions with larger acyl groups. Full kinetics could not be determined with 4-nitrophenyl octanoate due to solubility issues, so it is possible that Cqestβ2[1] has both a greater turnover number (k_{cat}) and Michaelis–Menten constant (K_M) with this substrate than the others. Thus, Cqestβ2[1] demonstrates good specificity with a range of acyl-chain lengths.

Table 2. Kinetic Parameters of Cqestβ2[1] with 4-Nitrophenyl Esters[b]

acyl group	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M (M^{-1} s^{-1})
acetyl	632.9 ± 5.7	74.9 ± 4.1	$1.41 \times 10^5 \pm 0.06 \times 10^5$
butyryl	891.1 ± 66.7	11.4 ± 2.5	$1.34 \times 10^6 \pm 0.22 \times 10^6$
octanoyl	N.D	N.D	$3.30 \times 10^4 \pm 0.12 \times 10^4$[a]

[a]Apparent k_{cat}/K_M, see Materials and Methods. [b]Values are the mean ± sample standard deviation (n = 3).

Interaction of Cqestβ2[1] with Insecticides. Enzyme reactivation experiments have previously suggested that Cqestβ2[1] is only able to slowly turnover OPs.[12,77] This is consistent with the proposal that Cqestβ2[1] confers metabolic resistance by sequestering the OP rather than the catalytic hydrolysis mechanism of resistance seen in LcαE7, where OPs are catalytically detoxified (Figure S1).[67] This mechanism was acquired from a single mutation in the glycine oxyanion hole of LcαE7, G137D, which has been proposed to act as a general base, activating a water molecule for the dephosphorylation reaction.[40,78,79] Molecular dynamic simulations demonstrated that the position and steric bulk of F309 were critical in allowing D137 to adopt catalytically productive conformations.[40] Both the expanded pocket of Cqestβ2[1] and the lack of an equivalent residue to F309 in LcαE7 would reduce the frequency of productive conformations of the equivalent G137D mutation in Cqestβ2[1] (Figure 3). This would dramatically reduce its viability and explains the observed prevalence of the sequestration mechanism in culicine mosquitoes.

To investigate the sequestration mechanism further, we studied the activity of Cqestβ2[1] directly with the model OP, diethylumbelliferyl phosphate (DEUP), using stopped flow kinetics (Figure 4 and Table S1 for chemical structure). Hydrolysis of DEUP results in the formation of 7-hydroxy-4-methyl coumarin (HMC), which can be monitored fluorometrically. This showed that the interaction of Cqestβ2[1] with DEUP involves a presteady state burst phase, where phosphorylation releases HMC, followed by very slow turnover (dephosphorylation) (Figure 4A). In comparison to LcαE7 wild-type, both the rate of the pre-steady state burst phase (3.0 vs 1.8 s^{-1} for LcαE7) and the steady state turnover (0.7 s^{-1} vs 1.0 s^{-1} for LcαE7) are comparable. However, the K_M associated with the burst phase is significantly higher for Cqestβ2[1]. The steady state turnover of Cqestβ2[1] is much slower than the G137D LcαE7 mutant associated with resistance in *L. cuprina*

DOI: 10.1021/acs.biochem.7b00774
Biochemistry 2017, 56, 5512–5525

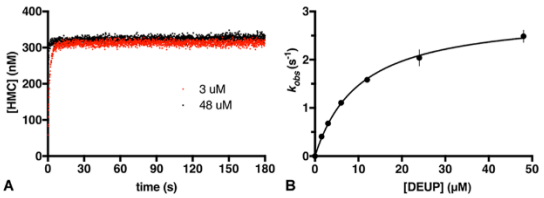


Figure 4. (A) Representative curves for Cqest$\beta 2^1$ hydrolysis of two concentrations of DEUP monitored through the formation of HMC using a stopped-flow fluorometric assay at pH 7.5. (B) Curve of Cqest$\beta 2^1$ presteady state burst phase kinetics showing the rate of the pre-steady state burst phase determined over a range of DEUP concentrations. Each point has three replicates and error bars are shown as lines.

(20-fold) (Table 3, Figure 4B).[40] This confirms that the role of Cqest$\beta 2^1$ in insecticide resistance is through sequestration.

Table 3. Stopped-Flow Kinetic Parameters for the Burst Phase Hydrolysis of DEUP by Cqest$\beta 2^1$ Compared to Previous Data Determined for LcαE7 by Mabbitt, et al. (2016)[40b]

	k_3 (s^{-1})	K_M (μM)[a]	k_s(s^{-1}) × 10^{-4}
Cqest$\beta 2^1$	3.0 ± 0.1	9.7 ± 0.7	0.7 ± 0.2
LcαE7 wild-type	1.8 ± 0.2	<1.5	1.0 ± 0.1
LcαE7 G137D	1.3 ± 0.1	8.6 ± 1.3	14.0 ± 0.2

[a]Where $K_M = (k_3 + k_2)/k_1$. [b]Values are the mean ± sample standard deviation (n = 3).

In *C. quinquefasciatus*, elevated CBE expression has been found to correlate with resistance to a wide range of insecticides.[8] To test whether these biological observations are consistent with the *in vitro* interactions between Cqest$\beta 2^1$ and the insecticides, we measured its inhibition with representatives of three major insecticide classes: OPs (paraoxon-ethyl and temephos), carbamates (propoxur), and SPs (fluorogenic SP analogues designed to resemble all eight cypermethrin diastereomers[41]) (Table 4 and Table S1 for chemical structures). We observed a strong interaction with the OPs paraoxon-ethyl and temephos (K_d from 66 nM to 70 μM). It has previously been found that the k_i for susceptible *C. pipiens* AChE with paraoxon-ethyl is 1.6 × 10^5 M^{-1} min^{-1},[80] which suggests that Cqest$\beta 2^1$ (k_i = 4.0 × 10^7 M^{-1} min^{-1}) interacts with paraoxon-ethyl with substantially higher sensitivity than does its AChE target, consistent with it being an effective sequestration CBE. While Cqest$\beta 2^1$ still shows good affinity with the carbamate propoxur (K_d = 614 μM), it is significantly lower than with the OPs. A previous study found that the k_i for *C. pipiens* AChE with propoxur is 1.45 × 10^5 M^{-1} min^{-1}, around 100-fold more sensitive than Cqest$\beta 2^1$ is to propoxur.[80] This suggests that the overexpression of Cqest$\beta 2^1$ might confer some protection against carbamates in the field, but not as effectively as for OPs, which is broadly consistent with previous studies.[12,81−83] The assays of Cqest$\beta 2^1$ with the cypermethrin analogues showed neither activity nor inhibition. This suggests that alternate mechanisms, not directly involving Cqest$\beta 2^1$, may be involved in SP resistance in *C. quinquefasciatus*.

Investigating the Sequence Diversity of Cqest$\beta 2^1$ Isoforms through a Comparison with Cqest$\beta 1$. There are as many as 11 naturally occurring isoforms of Cqest$\beta 2^1$.[15] It is not yet clear how these isoforms differ in their reactivity with insecticides, stability, and/or optimization for expression in *Culex* mosquitoes. BLAST was used to collect the Estβ CBE isoforms from *C. pipiens* and *C. quinquefasciatus* (the two most widespread *Culex* species).[84] By aligning these sequences it was found that there are only 41 sites that differ from the Cqest$\beta 2^1$ amino acid sequence (Figures S1). The majority of these sites are shared among the different isoforms. The mutations at 25 of them were found to have significant changes to either the residue's polarity, functionality, and/or size. Mapping the changes onto the structure of Cqest$\beta 2^1$ showed that none of them occur in the active site, and only five are near it, three in the second shell and two as a part of its entrance (Figure 5A).

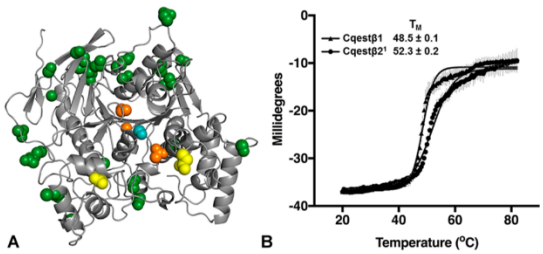


Figure 5. (A) A cartoon representation of the structure of Cqest$\beta 2^1$ looking into the active site pocket and showing the sites of amino acid mutations found in Cqest$\beta 2^1$ isoforms. The catalytic serine is shown in light blue, sites on the exterior of the structure in green, sites in the interior in orange and on the entrance tunnel in yellow. (B) Thermal denaturation curves of Cqest$\beta 2^1$ and Cqest$\beta 1$ determined using circular dichroism displaying calculated melting temperatures (T_M). Error bars are show as gray lines.

To investigate the effect of these mutations we selected the most divergent isoform, Cqest$\beta 1$.[27] Cqest$\beta 1$ includes all of the most significant mutations near the active site in the Cqest$\beta 2^1$ isoforms and has a total of 16 mutations from Cqest$\beta 2^1$ (Figure 5A). Uniquely, Cqest$\beta 1$ is not always coamplified with an Estα but confers resistance solely from its own amplification.[26] To compare Cqest$\beta 1$ with Cqest$\beta 2^1$ we expressed and purified the

Table 4. Inhibition Constants of Cqest$\beta 2^1$ with a Range of Insecticides[a]

inhibitor	CBE	k_2 (min^{-1})	K_d (μM)[b]	k_i(M^{-1} min^{-1})[b]
paraoxon-ethyl	Cqest$\beta 2^1$	2.67 ± 0.20	0.0663 ± 0.0056	4.03 × 10^7 ± 0.14 × 10^7
	Cqest$\beta 1$	6.96 ± 2.19	3.09 ± 1.08	2.27 × 10^6 ± 0.11 × 10^6
Propoxur	Cqest$\beta 2^1$	0.954 ± 0.210	614 ± 89	1.51 × 10^3 ± 0.24 × 10^3
	Cqest$\beta 1$	0.840 ± 0.060	593.0 ± 93.4	1.54 × 10^3 ± 0.14 × 10^3
Temephos	Cqest$\beta 2^1$	0.804 ± 0.432	69.7 ± 27.4	1.17 × 10^4 ± 0.16 × 10^4
	Cqest$\beta 1$	0.768 ± 0.066	83.95 ± 11.90	9.16 × 10^3 ± 0.54 × 10^3

[a]Values are the mean ± sample standard deviation (n = 3). [b]Where $K_d = (k_{-1}/k_1)$ and $k_i = k_2/(K_d)$.

DOI: 10.1021/acs.biochem.7b00774
Biochemistry 2017, 56, 5512−5525

enzyme in *E. coli*. Interestingly, Cqestβ1 has one mutation, A213R, which results in complete loss of soluble expression. As all other isoforms have alanine in this position, this mutation was excluded (Figure S2). Inspection of the structure shows the pocket containing A213 could not accommodate an arginine without significant disruption to the protein structure. This raises the possibility that the identification of this mutation could have been a sequencing error.[27] To determine whether there is a significant difference in thermostability between these enzymes, we conducted thermal denaturation experiments monitored through circular dichroism (CD) (Figure 5B). The melting temperature of Cqestβ2[1] was significantly higher (a difference of 3.9 °C) demonstrating greater thermostability. This is suggestive of greater stability and could be a significant factor in the wider distribution of Cqestβ2[1] over Cqestβ1.

To determine whether the two isoforms differ in their interaction with insecticides, we measured the reactivity of Cqestβ1 with the same range of insecticides as Cqestβ2[1] (Table 4). This demonstrated that it reacts with all but one of the insecticides tested with a similar sensitivity to Cqestβ2[1], including the lack of activity with the fluorescent cypermethrin analogues (Figure 6). The only interaction that differed was with paraoxon-ethyl, where Cqestβ1 had a $K_d \approx$ 50-fold lower and a $k_i \approx$ 20-fold less sensitivity, demonstrating that it too would be an effective sequestration CBE (Table 4). This suggests that the mutations that differ between these isoforms have not caused a major change in their sensitivity for these insecticides. Unless there are major epistatic effects in some combinations of these mutations, it seems likely that all the other isoforms will have similar interactions with insecticides. The similarity in the interaction of Cqestβ2[1] and Cqestβ1 with insecticides also suggests that future targeted design of inhibitors based on the structure of Cqestβ2[1] is likely to be broadly applicable across the isoforms.

■ SUMMARY

In this work we have provided new molecular-level understanding of the function of the primary insecticide-sequestering protein of the mosquito disease vector *Culex quinquefasciatus*. Our results reveal structural similarity to the target of organophosphate and carbamate insecticides, AChE, consistent with Cqestβ2[1] acting as a decoy "sponge" to intercept insecticide before they can reach their target. SSNs reveal that CBEs are a more diverse family than previously thought, with a number of subfamilies, several of which are now functionally annotated. Notably, the CBEs that have been recruited through evolution to an insecticide sequestration/detoxification role appear to predominantly come from a small number of these subfamilies involved in metabolism, with CBEs involved in hormone/odorant/semiochemical/neurotransmitter functions being generally excluded. Structural analysis reveals that insecticide-sequestering CBEs typically have much larger and less specialized substrate binding pockets than structures of hormone (JHE), odorant (Est6), and neurotransmitter (AChE) specific CBEs. We show, using stopped-flow kinetic analysis, that binding of OPs occurs rapidly and with high affinity, after which a covalent intermediate is formed that is essentially irreversible and leads to prolonged inhibition of the enzyme. Finally, we studied the effect of the substantial sequence diversity within isoforms of the Cqestβ proteins, which revealed this sequence variation is mostly neutral with respect to the binding of OPs. This structure and the new insight into the mechanism of insecticide binding should provide a basis for future studies aimed at combating insecticide resistance through the development of synergists to knockout the insecticide-sequestration function of these CBEs, and the development of more judicious approaches to the application of insecticides.

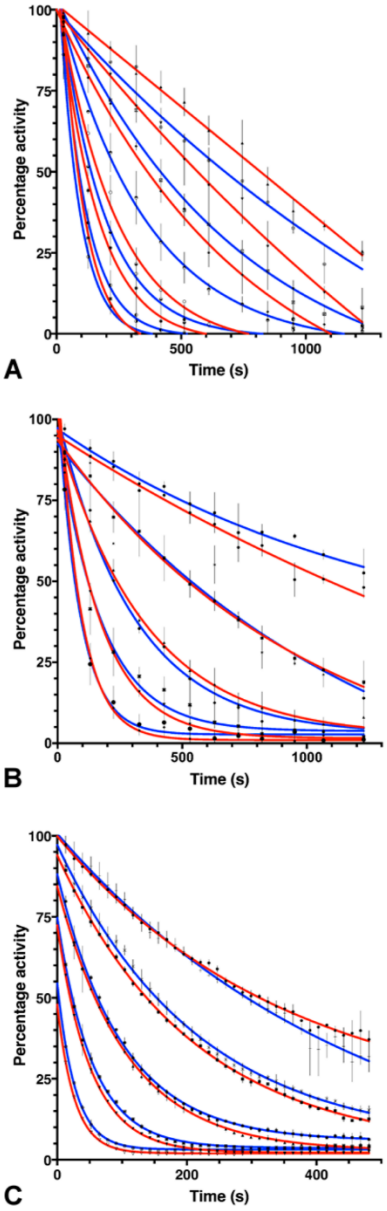


Figure 6. Percentage activity of Cqestβ2[1] (blue) and Cqestβ1 (red) after varying incubation time with insecticides. (A) For propoxur a 2-fold serial dilution of concentrations from 1600 μM to 50 μM was used. (B) For temephos a 2-fold serial dilution of concentrations from 60 μM to 3.75 μM was used. (C) For paraoxon, inhibition was determined in the presence of substrate with no incubation time using 2-fold serial dilution of concentrations from 9600 nM to 600 nM.

DOI: 10.1021/acs.biochem.7b00774
Biochemistry 2017, 56, 5512−5525

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.7b00774.

A reaction scheme for LcαE7 G137D, Cqestβ2[1], and AChE with OP insecticide showing the different paths, a list of NCBI Fasta files included in the SSN to aid group classification, a list of the UniProt IDs of insecticide resistance CBEs used in the SSNs, a table of the chemical structures mentioned in the results and discussion, a multiple sequence alignment of B2 isoforms from the *Culex* assemblage including a list of Genbank accession numbers (PDF)

Accession Codes

PDB ID: 5W1U; PDB ID: 4FG5; PDB ID: 1QO9; PDB ID: 4FNM; PDB ID: 2FJ0; PDB ID: 5THM; Genbank accession number: CAA83643; Genbank accession number: AAA28289.1.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: colin.jackson@anu.edu.au.

ORCID

Colin J. Jackson: 0000-0001-6150-3822

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This research was undertaken on the MX2 beamline at the Australian Synchrotron, part of ANSTO. This research was supported by the Australian Research Council (Future Fellowship to C.J.J.; FT140101059), Australian Science and Industry Endowment Fund (C.J.J. and P.D.M.; PF14-099), and by an Australian Government Research Training Program (RTP) Scholarship (D.H.H.).

■ REFERENCES

(1) Oakeshott, J. G., Claudianos, C., Campbell, P. M., Newcomb, R. D., and Russell, R. J. (2005) Biochemical genetics and genomics of insect esterases, in *Comprehensive Molecular Insect Science*, pp 309−381, Elsevier Ltd.

(2) Roush, R. T., and Tabashnik, B. E. (1990) Pesticide resistance in arthropods (Roush, R. T., and Tabashnik, B. E., Eds.) Springer Science & Business Media, Boston, MA.

(3) Sun, L., Zhou, X., Zhang, J., and Gao, X. (2005) Polymorphisms in a carboxylesterase gene between organophosphate-resistant and -susceptible Aphis gossypii (Homoptera: Aphididae). *J. Econ. Entomol. 98*, 1325−1332.

(4) Devonshire, A. L., Field, L. M., Foster, S. P., Moores, G. D., Williamson, M. S., and Blackman, R. L. (1998) The evolution of insecticide resistance in the peach-potato aphid, Myzus persicae. *Philos. Trans. R. Soc., B 353*, 1677−1684.

(5) Parker, A. G., Russell, R. J., Oakeshott, J. G., and Delves, A. c. (1991) Biochemistry and physiology of esterases in organophosphate-susceptible and -resistant strains of the Australian sheep blowfly, Lucilia cuprina. *Pestic. Biochem. Physiol. 41*, 305−318.

(6) Marcombe, S., Mathieu, R. B., Pocquet, N., Riaz, M. A., Poupardin, R., Sélior, S., Darriet, F., Reynaud, S., Yébakima, A., Corbel, V., David, J. P., and Chandre, F. (2012) Insecticide resistance in the dengue vector aedes aegypti from martinique: Distribution, mechanisms and relations with environmental factors. *PLoS One 7*, e30989.

(7) Mouchès, C., Pasteur, N., Bergé, J. B., Hyrien, O., Raymond, M., de Saint Vincent, B. R., de Silvestri, M., and Georghiou, G. P. (1986) Amplification of an esterase gene is responsible for insecticide resistance in a California Culex mosquito. *Science 233*, 778−780.

(8) Liu, Y., Zhang, H., Qiao, C., Lu, X., and Cui, F. (2011) Correlation between carboxylesterase alleles and insecticide resistance in Culex pipiens complex from China. *Parasites Vectors 4*, 236.

(9) Ketterman, a J., Jayawardena, K. G., and Hemingway, J. (1992) Purification and characterization of a carboxylesterase involved in insecticide resistance from the mosquito Culex quinquefasciatus. *Biochem. J. 287*, 355−360.

(10) Aranda, J., Cerqueira, N. M. F. S. a, Fernandes, P. A., Roca, M., Tuñón, I., and Ramos, M. J. (2014) The catalytic mechanism of carboxylesterases: a computational study. *Biochemistry 53*, 5820−5829.

(11) Aldridge, W. N. (1950) Some properties of specific cholinesterase with particular reference to the mechanism of inhibition by diethyl p-nitrophenyl thiophosphate (E 605) and analogues. *Biochem. J. 46*, 451−460.

(12) Karunaratne, S. H., Jayawardena, K. G., Hemingway, J., and Ketterman, a J. (1993) Characterization of a B-type esterase involved in insecticide resistance from the mosquito Culex quinquefasciatus. *Biochem. J. 294*, 575−579.

(13) Jackson, C. J., Liu, J.-W., Carr, P. D., Younus, F., Coppin, C., Meirelles, T., Lethier, M., Pandey, G., Ollis, D. L., Russell, R. J., Weik, M., and Oakeshott, J. G. (2013) Structure and function of an insect α-carboxylesterase (αEsterase7) associated with insecticide resistance. *Proc. Natl. Acad. Sci. U. S. A. 110*, 10177−10182.

(14) Gordon, J. R., and Ottea, J. (2012) Association of esterases with insecticide resistance in Culex quinquefasciatus (Diptera: Culicidae). *J. Econ. Entomol. 105*, 971−978.

(15) Zhang, H. H., Meng, F., Qiao, C. C., and Cui, F. (2012) Identification of resistant carboxylesterase alleles in Culex pipiens complex via PCR-RFLP. *Parasites Vectors 5*, 209.

(16) Hemingway, J., and Karunaratne, S. H. (1998) Mosquito carboxylesterases: a review of the molecular biology and biochemistry of a major insecticide resistance mechanism. *Med. Vet. Entomol. 12*, 1−12.

(17) Vaughan, a, and Hemingway, J. (1995) Mosquito carboxylesterase Est alpha 2(1) (A2). Cloning and sequence of the full-length cDNA for a major insecticide resistance gene worldwide in the mosquito Culex quinquefasciatus. *J. Biol. Chem. 270*, 17044−17049.

(18) Labbe, P., Lenormand, T., and Raymond, M. (2005) On the worldwide spread of an insecticide resistance gene: a role for local selection. *J. Evol. Biol. 18*, 1471−1484.

(19) Pasteur, N., Nancé, E., and Bons, N. (2001) Tissue Localization of Overproduced Esterases in the Mosquito Culex pipiens (Diptera: Culicidae). *J. Med. Entomol. 38*, 791−801.

(20) Raymond, M., Chevillon, C., Guillemaud, T., Lenormand, T., and Pasteur, N. (1998) An overview of the evolution of overproduced esterases in the mosquito Culex pipiens. *Philos. Trans. R. Soc., B 353*, 1707−1711.

(21) Callaghan, A., Guillemaud, T., Makate, N., and Raymond, M. (1998) Polymorphisms and fluctuations in copy number of amplified esterase genes in Culex pipiens mosquitoes. *Insect Mol. Biol. 7*, 295−300.

(22) Vaughan, a, Hawkes, N., and Hemingway, J. (1997) Co-amplification explains linkage disequilibrium of two mosquito esterase genes in insecticide-resistant Culex quinquefasciatus. *Biochem. J. 325* (2), 359−365.

(23) Karunaratne, S. H. P. P., and Hemingway, J. (2001) Malathion resistance and prevalence of the malathion carboxylesterase mechanism in populations of mosquito vectors of disease in Sri Lanka. *Bull. World Health Organ. 79*, 1060−1064.

(24) Yan, S., Nan, P., Cui, F., Wu, Z., and Qiao, C. (2013) Distribution and dynamics of esterase alleles in Culex pipiens complex in China. *J. Asia-Pac. Entomol. 16*, 43−48.

(25) Sogorb, M. A., and Vilanova, E. (2002) Enzymes involved in the detoxification of organophosphorus, carbamate and pyrethroid insecticides through hydrolysis. *Toxicol. Lett. 128*, 215−228.

DOI: 10.1021/acs.biochem.7b00774
Biochemistry 2017, 56, 5512−5525

(26) Vaughan, A., Rodriguez, M., and Hemingway, J. (1995) The independent gene amplification of electrophoretically indistinguishable B esterases from the insecticide-resistant mosquito Culex quinquefasciatus. *Biochem. J. 305* (2), 651−658.
(27) Mouches, C., Pauplin, Y., Agarwal, M., Lemieux, L., Herzog, M., Abadon, M., Beyssat-Arnaouty, V., Hyrien, O., de Saint Vincent, B. R., and Georghiou, G. P. (1990) Characterization of amplification core and esterase B1 gene responsible for insecticide resistance in Culex. *Proc. Natl. Acad. Sci. U. S. A. 87*, 2574−2578.
(28) Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., and Smith, H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods 6*, 343−345.
(29) Nallamsetty, S., Kapust, R. B., Tözsér, J., Cherry, S., Tropea, J. E., Copeland, T. D., and Waugh, D. S. (2004) Efficient site-specific processing of fusion proteins by tobacco vein mottling virus protease in vivo and in vitro. *Protein Expression Purif. 38*, 108−115.
(30) Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., and Bairoch, A. (2003) ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res. 31*, 3784−3788.
(31) Walter, T. S., Meier, C., Assenberg, R., Au, K.-F., Ren, J., Verma, A., Nettleship, J. E., Owens, R. J., Stuart, D. I., and Grimes, J. M. (2006) Lysine methylation as a routine rescue strategy for protein crystallization. *Structure 14*, 1617−1622.
(32) McPhillips, T. M., McPhillips, S. E., Chiu, H.-J., Cohen, A. E., Deacon, A. M., Ellis, P. J., Garman, E., Gonzalez, A., Sauter, N. K., Phizackerley, R. P., Soltis, S. M., and Kuhn, P. (2002) Blu-Ice and the Distributed Control System: software for data acquisition and instrument control at macromolecular crystallography beamlines. *J. Synchrotron Radiat. 9*, 401−406.
(33) Leslie, A. G. W., and Powell, H. R. (2007) Processing diffraction data with MOSFLM, in *Evolving Methods for Macromolecular Crystallography* (Read, R. J., and Sussman, J. L., Eds.) pp 41−51, Springer Netherlands, Dordrecht.
(34) Vagin, A., and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement. *J. Appl. Crystallogr. 30*, 1022−1025.
(35) Langer, G., Cohen, S. X., Lamzin, V. S., and Perrakis, A. (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat. Protoc. 3*, 1171−1179.
(36) Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 60*, 2126−2132.
(37) Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 67*, 355−367.
(38) Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 67*, 235−242.
(39) Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 66*, 213−221.
(40) Mabbitt, P. D., Correy, G. J., Meirelles, T., Fraser, N. J., Coote, M. L., and Jackson, C. J. (2016) Conformational disorganization within the active site of a recently evolved organophosphate hydrolase limits its catalytic efficiency. *Biochemistry 55*, 1408−1417.
(41) Huang, H., Stok, J. E., Stoutamire, D. W., Gee, S. J., and Hammock, B. D. (2005) Development of optically pure pyrethroid-like fluorescent substrates for carboxylesterases. *Chem. Res. Toxicol. 18*, 516−527.
(42) Aldridge, W. N., and Reiner, E. (1972) Enzyme inhibitors as substrates: interactions of esterases with esters of organophosphorus and carbamic acids, in *Frontiers of Biology*, Vol. *26*, North-Holland Pub. Co., Amsterdam.
(43) Hart, G. J., and O'Brien, R. D. (1974) Stopped-flow studies of the inhibition of acetylcholinesterase by organophosphates in the presence of substrate. *Pestic. Biochem. Physiol. 4*, 239−244.
(44) Schrodinger LLC. (2010) *The PyMOL Molecular Graphics System*, Version 1.3r1.
(45) Bond, C. S. (2003) TopDraw: a sketchpad for protein structure topology cartoons. *Bioinformatics 19*, 311−312.
(46) Gerlt, J. A., Bouvier, J. T., Davidson, D. B., Imker, H. J., Sadkhin, B., Slater, D. R., and Whalen, K. L. (2015) Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A web tool for generating protein sequence similarity networks. *Biochim. Biophys. Acta, Proteins Proteomics 1854*, 1019−1037.
(47) Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res. 13*, 2498−2504.
(48) Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol. 372*, 774−797.
(49) Fraser, N. J., Liu, J.-W., Mabbitt, P. D., Correy, G. J., Coppin, C. W., Lethier, M., Perugini, M. A., Murphy, J. M., Oakeshott, J. G., Weik, M., and Jackson, C. J. (2016) Evolution of protein quaternary structure in response to selective pressure for increased thermostability. *J. Mol. Biol. 428*, 2359−2371.
(50) Nardini, M., and Dijkstra, B. W. (1999) α/β Hydrolase fold enzymes: the family keeps growing. *Curr. Opin. Struct. Biol. 9*, 732−737.
(51) Claudianos, C., Ranson, H., Johnson, R. M., Biswas, S., Schuler, M. A., Berenbaum, M. R., Feyereisen, R., and Oakeshott, J. G. (2006) A deficit of detoxification enzymes: pesticide sensitivity and environmental response in the honeybee. *Insect Mol. Biol. 15*, 615−636.
(52) Oakeshott, J. G., Johnson, R. M., Berenbaum, M. R., Ranson, H., Cristino, A. S., and Claudianos, C. (2010) Metabolic enzymes associated with xenobiotic and chemosensory responses in Nasonia vitripennis. *Insect Mol. Biol. 19*, 147−163.
(53) Sadd, B. M., Barribeau, S. M., Bloch, G., de Graaf, D. C., Dearden, P., Elsik, C. G., Gadau, J., Grimmelikhuijzen, C. J. P., Hasselmann, M., Lozier, J. D., Robertson, H. M., Smagghe, G., Stolle, E., Van Vaerenbergh, M., Waterhouse, R. M., Bornberg-Bauer, E., Klasberg, S., Bennett, A. K., Câmara, F., Guigó, R., Hoff, K., Mariotti, M., Munoz-Torres, M., Murphy, T., Santesmasses, D., Amdam, G. V., Beckers, M., Beye, M., Biewer, M., Bitondi, M. M. G., Blaxter, M. L., Bourke, A. F. G., Brown, M. J. F., Buechel, S. D., Cameron, R., Cappelle, K., Carolan, J. C., Christiaens, O., Ciborowski, K. L., Clarke, D. F., Colgan, T. J., Collins, D. H., Cridge, A. G., Dalmay, T., Dreier, S., du Plessis, L., Duncan, E., Erler, S., Evans, J., Falcon, T., Flores, K., Freitas, F. C. P., Fuchikawa, T., Gempe, T., Hartfelder, K., Hauser, F., Helbing, S., Humann, F. C., Irvine, F., Jermiin, L. S., Johnson, C. E., Johnson, R. M., Jones, A. K., Kadowaki, T., Kidner, J. H., Koch, V., Köhler, A., Kraus, F. B., Lattorff, H. M. G., Leask, M., Lockett, G. A., Mallon, E. B., Antonio, D. S. M., Marxer, M., Meeus, I., Moritz, R. F. A., Nair, A., Näpflin, K., Nissen, I., Niu, J., Nunes, F. M. F., Oakeshott, J. G., Osborne, A., Otte, M., Pinheiro, D. G., Rossié, N., Rueppell, O., Santos, C. G., Schmid-Hempel, R., Schmitt, B. D., Schulte, C., Simões, Z. L. P., Soares, M. P. M., Swevers, L., Winnebeck, E. C., Wolschin, F., Yu, N., Zdobnov, E. M., Aqrawi, P. K., Blankenburg, K. P., Coyle, M., Francisco, L., Hernandez, A. G., Holder, M., Hudson, M. E., Jackson, L., Jayaseelan, J., Joshi, V., Kovar, C., Lee, S. L., Mata, R., Mathew, T., Newsham, I. F., Ngo, R., Okwuonu, G., Pham, C., Pu, L.-L., Saada, N., Santibanez, J., Simmons, D., Thornton, R., Venkat, A., Walden, K. K. O., Wu, Y.-Q., Debyser, G., Devreese, B., Asher, C., Blommaert, J., Chipman, A. D., Chittka, L., Fouks, B., Liu, J., O'Neill, M. P., Sumner, S., Puiu, D., Qu, J., Salzberg, S. L., Scherer, S. E., Muzny, D. M., Richards, S., Robinson, G. E., Gibbs, R. A., Schmid-Hempel, P., and Worley, K. C. (2015) The genomes of two key bumblebee species with primitive eusocial organization. *Genome Biol. 16*, 76.

DOI: 10.1021/acs.biochem.7b00774
Biochemistry 2017, 56, 5512−5525

(54) Teese, M. G., Campbell, P. M., Scott, C., Gordon, K. H. J. J., Southon, A., Hovan, D., Robin, C., Russell, R. J., and Oakeshott, J. G. (2010) Gene identification and proteomic analysis of the esterases of the cotton bollworm, Helicoverpa armigera. *Insect Biochem. Mol. Biol. 40*, 1−16.
(55) Yu, Q.-Y., Lu, C., Li, W.-L., Xiang, Z.-H., and Zhang, Z. (2009) Annotation and expression of carboxylesterases in the silkworm, Bombyx mori. *BMC Genomics 10*, 553.
(56) Strode, C., Wondji, C. S., David, J.-P., Hawkes, N. J., Lumjuan, N., Nelson, D. R., Drane, D. R., Karunaratne, S. H. P. P., Hemingway, J., Black, W. C., and Ranson, H. (2008) Genomic analysis of detoxification genes in the mosquito Aedes aegypti. *Insect Biochem. Mol. Biol. 38*, 113−123.
(57) Lü, F.-G., Fu, K.-Y., Li, Q., Guo, W.-C., Ahmat, T., and Li, G.-Q. (2015) Identification of carboxylesterase genes and their expression profiles in the Colorado potato beetle Leptinotarsa decemlineata treated with fipronil and cyhalothrin. *Pestic. Biochem. Physiol. 122*, 86−95.
(58) Ramsey, J. S., Rider, D. S., Walsh, T. K., De Vos, M., Gordon, K. H. J., Ponnala, L., Macmil, S. L., Roe, B. A., and Jander, G. (2010) Comparative analysis of detoxification enzymes in Acyrthosiphon pisum and Myzus persicae. *Insect Mol. Biol. 19*, 155−164.
(59) Kanost, M. R., Arrese, E. L., Cao, X., Chen, Y.-R., Chellapilla, S., Goldsmith, M. R., Grosse-Wilde, E., Heckel, D. G., Herndon, N., Jiang, H., Papanicolaou, A., Qu, J., Soulages, J. L., Vogel, H., Walters, J., Waterhouse, R. M., Ahn, S.-J., Almeida, F. C., An, C., Aqrawi, P., Bretschneider, A., Bryant, W. B., Bucks, S., Chao, H., Chevignon, G., Christen, J. M., Clarke, D. F., Dittmer, N. T., Ferguson, L. C. F., Garavelou, S., Gordon, K. H. J., Gunaratna, R. T., Han, Y., Hauser, F., He, Y., Heidel-Fischer, H., Hirsh, A., Hu, Y., Jiang, H., Kalra, D., Klinner, C., König, C., Kovar, C., Kroll, A. R., Kuwar, S. S., Lee, S. L., Lehman, R., Li, K., Li, Z., Liang, H., Lovelace, S., Lu, Z., Mansfield, J. H., McCulloch, K. J., Mathew, T., Morton, B., Muzny, D. M., Neunemann, D., Ongeri, F., Pauchet, Y., Pu, L.-L., Pyrousis, I., Rao, X.-J., Redding, A., Roesel, C., Sanchez-Gracia, A., Schaack, S., Shukla, A., Tetreau, G., Wang, Y., Xiong, G.-H., Traut, W., Walsh, T. K., Worley, K. C., Wu, D., Wu, W., Wu, Y.-Q., Zhang, X., Zou, Z., Zucker, H., Briscoe, A. D., Burmester, T., Clem, R. J., Feyereisen, R., Grimmelikhuijzen, C. J. P., Hamodrakas, S. J., Hansson, B. S., Huguet, E., Jermiin, L. S., Lan, Q., Lehman, H. K., Lorenzen, M., Merzendorfer, H., Michalopoulos, I., Morton, D. B., Muthukrishnan, S., Oakeshott, J. G., Palmer, W., Park, Y., Passarelli, A. L., Rozas, J., Schwartz, L. M., Smith, W., Southgate, A., Vilcinskas, A., Vogt, R., Wang, P., Werren, J., Yu, X.-Q., Zhou, J.-J., Brown, S. J., Scherer, S. E., Richards, S., and Blissard, G. W. (2016) Multifaceted biological insights from a draft genome sequence of the tobacco hornworm moth, Manduca sexta. *Insect Biochem. Mol. Biol. 76*, 118−147.
(60) Pearce, S. L., Clarke, D. F., East, P. D., Elfekih, S., Gordon, K. H. J., Jermiin, L. S., McGaughran, A., Oakeshott, J. G., Papanikolaou, A., and Perera, O. P. (2017) Genomic innovations, transcriptional plasticity and gene loss underlying the evolution and divergence of two highly polyphagous and invasive Helicoverpa pest species. *BMC Biol. 15*, 63.
(61) Ahmed, F. H., Carr, P. D., Lee, B. M., Afriat-Jurnou, L., Mohamed, A. E., Hong, N.-S., Flanagan, J., Taylor, M. C., Greening, C., and Jackson, C. J. (2015) Sequence-structure-function classification of a catalytically diverse oxidoreductase superfamily in Mycobacteria. *J. Mol. Biol. 427*, 3554−3571.
(62) Brown, S. D., and Babbitt, P. C. (2012) Inference of functional properties from large-scale analysis of enzyme superfamilies. *J. Biol. Chem. 287*, 35−42.
(63) Atkinson, H. J., Morris, J. H., Ferrin, T. E., and Babbitt, P. C. (2009) Using sequence similarity networks for visualization of relationships across diverse protein superfamilies. *PLoS One 4*, e4345.
(64) Wogulis, M., Wheelock, C. E., Kamita, S. G., Hinton, A. C., Whetstone, P. A., Hammock, B. D., and Wilson, D. K. (2006) Structural studies of a potent insect maturation inhibitor bound to the juvenile hormone esterase of Manduca sexta. *Biochemistry 45*, 4045−4057.
(65) Younus, F., Fraser, N. J., Coppin, C. W., Liu, J.-W., Correy, G. J., Chertemps, T., Pandey, G., Maïbèche, M., Jackson, C. J., and Oakeshott, J. G. (2017) Molecular basis for the behavioral effects of the odorant degrading enzyme Esterase 6 in Drosophila. *Sci. Rep. 7*, 46188.
(66) Cui, F., Lin, Z., Wang, H., Liu, S., Chang, H., Reeck, G., Qiao, C., Raymond, M., and Kang, L. (2011) Two single mutations commonly cause qualitative change of nonspecific carboxylesterases in insects. *Insect Biochem. Mol. Biol. 41*, 1−8.
(67) Newcomb, R. D., Campbell, P. M., Ollis, D. L., Cheah, E., Russell, R. J., and Oakeshott, J. G. (1997) A single amino acid substitution converts a carboxylesterase to an organophosphorus hydrolase and confers insecticide resistance on a blowfly. *Proc. Natl. Acad. Sci. U. S. A. 94*, 7464−7468.
(68) Durand, N., Carot-Sans, G., Chertemps, T., Montagné, N., Jacquin-Joly, E., Debernard, S., and Maïbèche-Coisne, M. (2010) A diversity of putative carboxylesterases are expressed in the antennae of the noctuid moth Spodoptera littoralis. *Insect Mol. Biol. 19*, 87−97.
(69) Younus, F., Chertemps, T., Pearce, S. L., Pandey, G., Bozzolan, F., Coppin, C. W., Russell, R. J., Maïbèche-Coisne, M., and Oakeshott, J. G. (2014) Identification of candidate odorant degrading gene/enzyme systems in the antennal transcriptome of Drosophila melanogaster. *Insect Biochem. Mol. Biol. 53*, 30−43.
(70) Han, Q., Wong, D. M., Robinson, H., Ding, H., Lam, P. C. H., Totrov, M. M., Carlier, P. R., and Li, J. (2017) Crystal structure of acetylcholinesterase catalytic subunits of the malaria vector Anopheles gambiae. *Insect Sci. 00*, 1−4.
(71) Harel, M., Kryger, G., Rosenberry, T. L., Mallender, W. D., Lewis, T., Fletcher, R. J., Guss, J. M., Silman, I., and Sussman, J. L. (2000) Three-dimensional structures of Drosophila melanogaster acetylcholinesterase and of its complexes with two potent inhibitors. *Protein Sci. 9*, 1063−1072.
(72) Krissinel, E., and Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 60*, 2256−2268.
(73) Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y., and Liang, J. (2006) CASTp: Computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res. 34*, W116−W118.
(74) Birner-Gruenberger, R., Bickmeyer, I., Lange, J., Hehlert, P., Hermetter, A., Kollroser, M., Rechberger, G. N., and Kühnlein, R. P. (2012) Functional fat body proteomics and gene targeting reveal in vivo functions of Drosophila melanogaster α-Esterase-7. *Insect Biochem. Mol. Biol. 42*, 220−229.
(75) Rauwerdink, A., and Kazlauskas, R. J. (2015) How the same core catalytic machinery catalyzes 17 different reactions: the serine-histidine-aspartate catalytic triad of α/β-hydrolase fold enzymes. *ACS Catal. 5*, 6153−6176.
(76) Shafferman, A., Ordentlich, A., Barak, D., Stein, D., Ariel, N., and Velan, B. (1996) Aging of phosphylated human acetylcholinesterase: catalytic processes mediated by aromatic and polar residues of the active centre. *Biochem. J. 318* (3), 833−840.
(77) Jayawardena, K. G. I., Karunaratne, S. H. P. P., Ketterman, A. J., and Hemingway, J. (1994) Determination of the role of elevated B2 esterase in insecticide resistance in Culex quinquefasciatus (Diptera: Culicidae) from studies on the purified enzyme. *Bull. Entomol. Res. 84*, 39−43.
(78) Heidari, R., Devonshire, A. L., Campbell, B. E., Bell, K. L., Dorrian, S. J., Oakeshott, J. G., and Russell, R. J. (2004) Hydrolysis of organophosphorus insecticides by in vitro modified carboxylesterase E3 from Lucilia cuprina. *Insect Biochem. Mol. Biol. 34*, 353−363.
(79) Devonshire, A. L., Heidari, R., Bell, K. L., Campbell, P. M., Campbell, B. E., Odgers, W. A., Oakeshott, J. G., and Russell, R. J. (2003) Kinetic efficiency of mutant carboxylesterases implicated in organophosphate insecticide resistance. *Pestic. Biochem. Physiol. 76*, 1−13.
(80) Bourguet, D., Raymond, M., Fournier, D., Malcolm, C. A., Toutant, J. P., and Arpagaus, M. (1996) Existence of two

DOI: 10.1021/acs.biochem.7b00774
Biochemistry 2017, 56, 5512−5525

acetylcholinesterases in the mosquito Culex pipiens (Diptera:Culicidae). *J. Neurochem. 67*, 2115−2123.
(81) Li, C. X., Dong, Y. D., Song, F. L., Zhang, X. L., Gu, W. D., and Zhao, T. Y. (2009) Company amplification of estalpha2/estbeta2 and correlation between esterase gene copy number and resistance to insecticides in the field Culex pipiens pallens strains collected from Beijing, China. *J. Med. Entomol. 46*, 539−545.
(82) Peiris, H. T. R., and Hemingway, J. (1993) Characterization and inheritance of elevated esterases in organophosphorus and carbamate insecticide resistant Culex quinquefasciatus (Diptera: Culicidae) from Sri Lanka. *Bull. Entomol. Res. 83*, 127−132.
(83) Bisset, J. A., Rodriguez, M. M., Diaz, C., Ortiz, E., Marquetti, M. C., and Hemingway, J. (1990) The mechanisms of organophosphate and carbamate resistance in Culex quinquefasciatus (Diptera: Culicidae) from Cuba. *Bull. Entomol. Res. 80*, 245−250.
(84) Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol. 215*, 403−410.

DOI: 10.1021/acs.biochem.7b00774
Biochemistry 2017, 56, 5512−5525

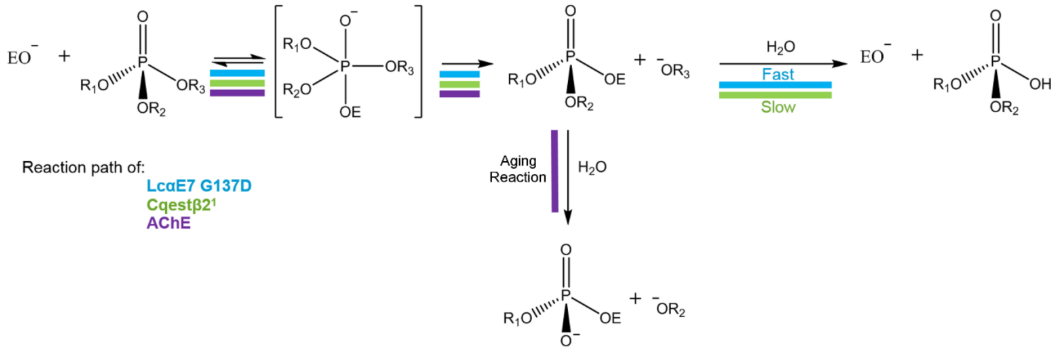
Structure of an Insecticide Sequestering Carboxylesterase from the Disease Vector *Culex quinquefasciatus:* What Makes an Enzyme a Good Insecticide Sponge? – Supporting information

Davis H. Hopkins[1] Nicholas J. Fraser[1] Peter D. Mabbitt[1] Paul D. Carr[1] John G. Oakeshott[2] Colin J. Jackson[1]*

[1]Research School of Chemistry, Australian National University, Canberra, ACT, 0200, Australia

[2]CSIRO, GPO Box 1700, Canberra, ACT, 2601, Australia

Figures S1. Reaction scheme for LcαE7 G137D, Cqestβ2[1] and AChE reacting with OP insecticide showing the different reaction pathways as colored lines where Fast and Slow are relative to each other.



List S1. NCBI Fasta Files for Sequences of Interest

>gi|3153851|gb|AAC36246.1| carboxylesterase [Anisopteromalus calandrae]
MERPEVKTLSGQVRGLKQISVEGIGFYAFKGIPYAKPPVGELRFKDPVPIEPWQEVREATEFGPMAAQ
FDVISKFSGGSDDCLYINVYTKKINSNVKQPVMFYIHGGGFIFGSGNDFFYGPDFLMRKDIVLVTFNY
RLGVFGFLNLEHEVAPGNQGLKDQVMALKWVRDNIANFGGDSENVTIFGESAGGASVHYLTVSPLAKG
LFHKAISQSGVFMNPGASVSGEPRKKAYELCELLGKKTTDPVEIVKFLRTVDTMKLIEHQGELQIQEL
QKKCLSAFVPGVDDKSPNPFMPFSREVAVEQAAHVPYLIGYNDREGTLLYKIFENDDFESKNLRFEEF
IHPNFAETLKRKKISLEDLKRMYFKNKKISKETTGKFIDLFSDMYFIQGIHQVARVQAERNSAPTYMY
QFTYDQGPNFSKGMFSIDEPGSTHMDELIYLFSMKFQETLNMEPIDKKSPHFRVMEQMVELWTNFAKY
GRPIPAPTELLPVHWLPMNDGTVLRYLNIGEELRMEKVLNIEERYDYKLICHREKV

>gi|17646748|gb|AAL41023.1|AF448479_1 juvenile hormone esterase [Tenebrio molitor]
MKLNVFVLISLHICFTLESVSTYAPKSPIVYTKYGPAIGLTEVSRSKREFMSFRGIPYAKPPVGELRF
AAPEPPEPWNFAINATQNGPICIQKNYFFSDPKVEGTEDCLYLNVYVPKVKGTALLPVMVFIHWGGFL
AGRGTSDYFGPEYIMDKDVILVTFNYRLGVFGFFTTLDDFAPGNYGLKDQVMALKFVQETIECFGGDK
NRVTIFGQSAGGGSVSLHLVSPLSRGLFQQAISQSGVALDLWAKPLNPLQVQITAALATFAGCQAQLA
NRADLLKCLREVDAQKLAETADKFKYFSIEPLTPYTLVTENKTAANPNPFLEKQPLSYLQDEEFLKVP
WIVGNVQDEGILRVAQLVRQPETLKALNDNFKTLITQLLALQVSVGENTTSLYDNMINFYLGGKSSID
VNDPKSVQGFIDLYTDRSFTYGTYQTAILQSQKGHTPIWMYNFDYRGQYSYGDQFAATTDDINFNWGV
SHCDDLLYLFKSPALFPDLEKESDLKMSEALVTMWTNFAIHGTPDPFNSLQWNSLNFNNTEEIKGSNL
KFLNMTGSSKTKTIEFKLQRGFYPERMQFWANQNLAENFPGLG

>gi|397511|emb|CAA52648.1| carboxylesterase E4 [Myzus persicae]
MKNTCGILLNLFLFIGCFLTCSASNTPKVQVHSGEIAGGFEYTYNGRKIYSFLGIPYASPPVQNNRFK
EPQPVQPWLGVWNATVPGSACLGIEFGSGSKIIGQEDCLFLNVYTPKLPQENSAGDLMNVIVHIHGGG
YYFGEGILYGPHYLLDNNDFVYVSINYRLGVLGFASTGDGVLTGNNGLKDQVAALKWIQQNIVAFGGD
PNSVTITGMSAGASSVHNHLISPMSKGLFNRAIIQSGSAFCHWSTAENVAQKTKYIANLMGCPTNNSV
EIVECLRSRPAKAIAKSYLNFMPWRNFPFTPFGPTVEVAGYEKFLPDIPEKLVPHDIPVLISIAQDEG
LIFSTFLGLENGFNELNNNWNEHLPHILDYNYTISNENLRFKTAQDIKEFYFGDKPISKETKSNLSKM
ISDRSFGYGTSKAAQHIAAKNTAPVYFYEFGYSGNYSYVAFFDPKSYSRGSSPTHGDETSYVLKMDGF
YVYDNEEDRKMIKTMVNIWATFIKSGVPDTENSEIWLPVSKNLADPFRFTKITQQQTFEAREQSTTGI
MNFGVAYH

>gi|397513|emb|CAA52649.1| carboxylesterase FE4 [Myzus persicae]
MKNTCGILLNLFLFIGCFLTCSASNTPKVQVHSGEIAGGFEYTYNGRKIYSFLGIPYASPPVQNNRFK
EPQPVQPWLGVWNATVPGSACLGIEFGSGSKIIGQEDCLFLNVYTPKLPQENSAGDLMNVIVHIHGGG
YYFGEGILYGPHYLLDNNDFVYVSINYRLGVLGFASTGDGVLPGNNGLKDQVAALKWIQQNIVAFGGD
PNSVTITGMSAGASSVHNHLISPMSKGLFNRAIIQSGSAFCHWSTAENVAQKTKYIANLLGCPTNNSV
EIVECLRSRPAKAIAKSYLNFMPWRNFPFTPFGPTVEVAGYEKFLPDIPEKLVPHDIPVLISIAQDEG
LIFSTFLGLENGFNELNNNWNEHLPHILDYNYTISNENLRFKTAQDIKEFYFGDKPISKETKSNLSKM
ISDRSFGYGTSKAAQHIAAKNTAPVYFYEFGYSGNYSYVAFFDPKSYSRGSSPTHGDETNYVLKVDGF
TVYDNEEDRKMIKTMVNIWATFIKSGVPDTENSEIWLPVSKNPADLFRFTKITQQQTFEAREQSTMAI
MNFGVAYHYQNILNLMCQMT

>gi|475068|emb|CAA83643.1| serine esterase [Culex quinquefasciatus B2]
MSLESLTVQTKYGPVRGKRSVSLLGQEYVSFQGIPYARAPEGELRFKAPVPPQNWTETLDCSQQCEPC
YHFDRRLQKIVGCEDSLKINVFAKEINPSKPLPVMLYIYGGGFTEGTSGTELYGPDFLVQKDIVLVSF
NYRIGALGFLCCQSEQDGVPGNAGLKDQNLAIRWVLENIAAFGGDPKRVTLVGHSAGAASVQYHLISD
ASKDLFQRAIVMSGSTYNSWSLTRQRNWVEKLAKAIGWDGQGGESGALRFLKAAKPEDIVANQEKLLT
DQDMQDDIFTPFGPTVEPYLTEQCMIPKEPFEMARTAWGDKIDIMIGGTSEEGLLLLQKIKLQPELLS
HPHLFLGNVPPNLKISMEKRIEFAAKLKQRYYPDSSPSMENNLGYVHMMSDRVFWHGLHRTILARAAR
SRARTFVYRICLDSEFYNHYRIMMIDPKLRGTAHADELSYLFSNFTQQVPGKETFEYRGLQTLVDVFT
AFVINGDPNCGMTAKSGVVFEPNAQTKPTFKCLNIANDGVAFVDYPDADRLDMWDAMYVNDELF

>gi|28574499|ref|NP_788500.1| esterase 6, isoform A [Drosophila melanogaster]
MNYVGLGLIIVLSCLWLGSNASDTDDPLLVQLPQGKLRGRDNGSYYSYESIPYAEPPTGDLRFEAPEP
YKQKWSDIFDATKTPVACLQWDQFTPGANKLVGEEDCLTVSVYKPKNSKRNSFPVVAHIHGGAFMFGA
AWQNGHENVMREGKFILVKISYRLGPLGFVSTGDRDLPGNYGLKDQRLALKWIKQNIASFGGEPQNVL
LVGHSAGGASVHLQMLREDFGQLARAAFSFSGNALDPWVIQKGARGRAFELGRNVGCESAEDSTSLKK
CLKSKPASELVTAVRKFLIFSYVPFAPFSPVLEPSDAPDAIITQDPRDVIKSGKFGQVPWAVSYVTED
GGYNAALLLKERKSGIVIDDLNERWLELAPYLLFYRDTKTKKDMDDYSRKIKQEYIGNQRFDIESYSE
LQRLFTDILFKNSTQESLDLHRKYGKSPAYAYVYDNPAEKGIAQVLANRTDYDFGTVHGDDYFLIFEN
FVRDVEMRPDEQIISRNFINMLADFASSDNGSLKYGECDFKDNVGSEKFQLLAIYIDGCQNRQHVEFP

>gi|109157641|pdb|2FJ0|A Chain A, Crystal Structure Of Juvenile Hormone Esterase From Manduca Sexta, With Otfp Covalently Attached
RIPSTEEVVVRTESGWIRGLKRRAEGNKSYASFRGVPYAKQPLGELRFKELQPLEPWQDELDATQEGP
VCQQTDVLYGRIMRPRGMSEACIHANIHVPYYALPRDAADKNRFAGLPVLVFIHGGGFAFGSGDSDLH
GPEYLVSKDVIVITFNYRLNVYGFLSLNSTSVPGNAGLRDMVTLLKWVQRNAHFFGGRPDDVTLMGQS
AGAAATHILSLSKAADGLFRRAILMSGTSSSAFFTTNPVFAQYINKLFVTNIGITATDPEEIHQKLIE
MPAEKLNEANRFLLEQFGLTTFFPVVESPINGVTTILDGDPEQLIAKGRGKHIPLIIGFTDAECEIFR
RQFEQIDIVSKIKENPGILVPLSVLFSSAPDTVAEITKAMHEKYFKKSVDMEGYIELCTDSYFMYPAI
SLAIKRARSNGAPVYLYQFSFDGDYSVFREVNHLNFEGAGHIEDLTYVFRTNSMLGGHASFPPHDKDD
HMKYWMTSFITNFMKYSNPVTDAKLWPEVRADNLRYQDIDTPDVYQNVKPHSEQRDMLDFFDSIYNWN
GTSYCIK

>gi|170039237|ref|XP_001847448.1| acetylcholinesterase 1 [Culex quinquefasciatus]
MEIRGLITRLLGPCHLRHLILCSLGLYSILVQSVHCRHHDIGSSVAHQLGSKYSQSSSLSSSSQSSSS
LAEEATLNKDSDAFFTPYIGHGDSVRIVDAELGTLEREHIHSTTTRRRGLTRRESSSDATDSDPLVIT
TDKGKIRGTTLEAPSGKKVDAWMGIPYAQPPLGPLRFRHPRPAERWTDRGHVFGDFPGATMWNRNTPL
SEDCLYINVFVPRPRPKNAAVMLWIFGGGFYSGTATLDVYDHRTLASEENVIVVSLQYRVASLGFLFL
GTPEAPGNAGLFDQNLALRWVRDNIHRFGGDPSRVTLFGESAGAVSVSLHLLSALSRDLFQRAILQSG
SPTAPWALVSREEATLRALRLAEAVNCPHDATKLSDAVECLRTKDPNELVDNEWGTLGICEFPFVPVV
DGAFLDETPQRSLASGRFKKTDILTGSNTEEGYYFIIYYLTELLRKEEGVTVTREEFLQAVRELNPYV
NGAARQAIVFEYTDWIEPDNPNSNRDALDKMVGDYHFTCNVNEFAQRYAEEGNNVFMYLYTHRSKGNP
WPRWTGVMHGDEINYVFGEPLNSALGYQDDEKDFSRKIMRYWSNFAKTGNPNPSTPSVDLPEWPKHTA
HGRHYLELGLNTTFVGRGPRLRQCAFWKKYLPQLVAATSNLQVTPAPSVPCESSSTSYRSTLLLIVTL
LLVTRFKI

List S2. UniProt ID of Insecticide Resistance CBEs Used in SSN Labelling

Q17B31_AEDAE, Q17B28_AEDAE, A0A0C4JZC5_CYDPO, B3SST0_BEMTA, O76177_APHGO, O76177_APHGO, EST6_APIME, T1VXB0_LOCMI, O61726_9HYME, O61727_9HYME, Q95UJ4_APHGO, O76177_APHGO, Q6RYX3_APHGO, W8EFX5_LOCMI, W8EC29_LOCMI, W8E8J7_LOCMI, A0A0H4D277_BACDO, A0A0H4DA18_BACDO, T1VXW3_LOCMI, ESTE_MYZPE, ESTF_MYZPE, Q25252_LUCCU, Q23734_CULQU.

Table S1. Chemical Structures Mentioned in the Results and Discussion

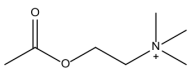
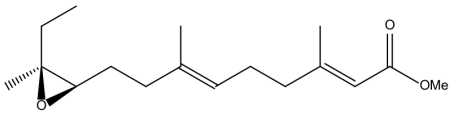
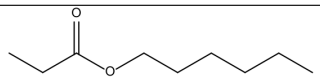
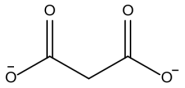
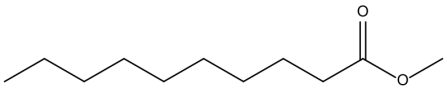
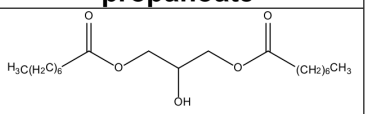
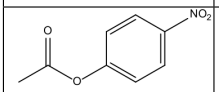
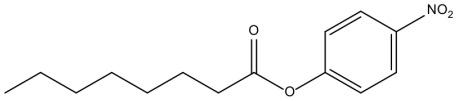
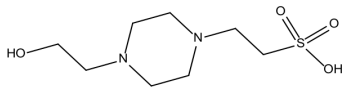
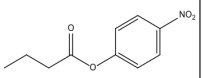
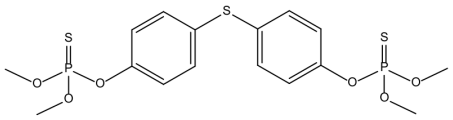
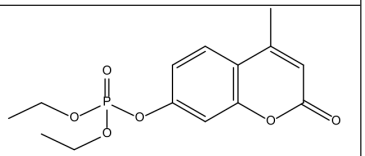
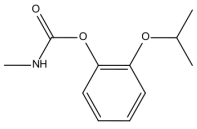
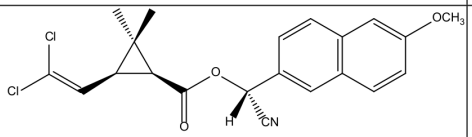
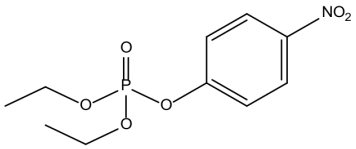
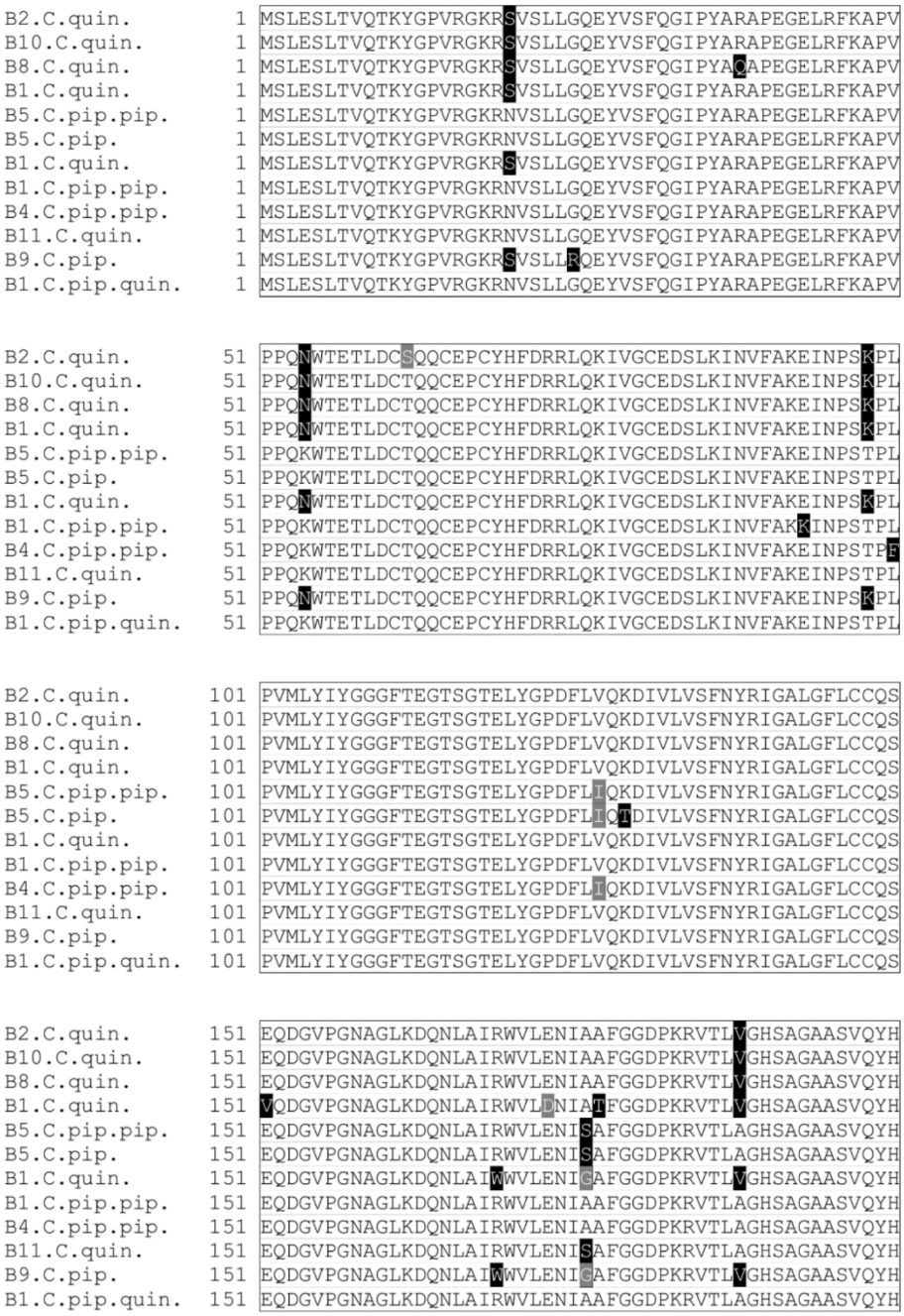
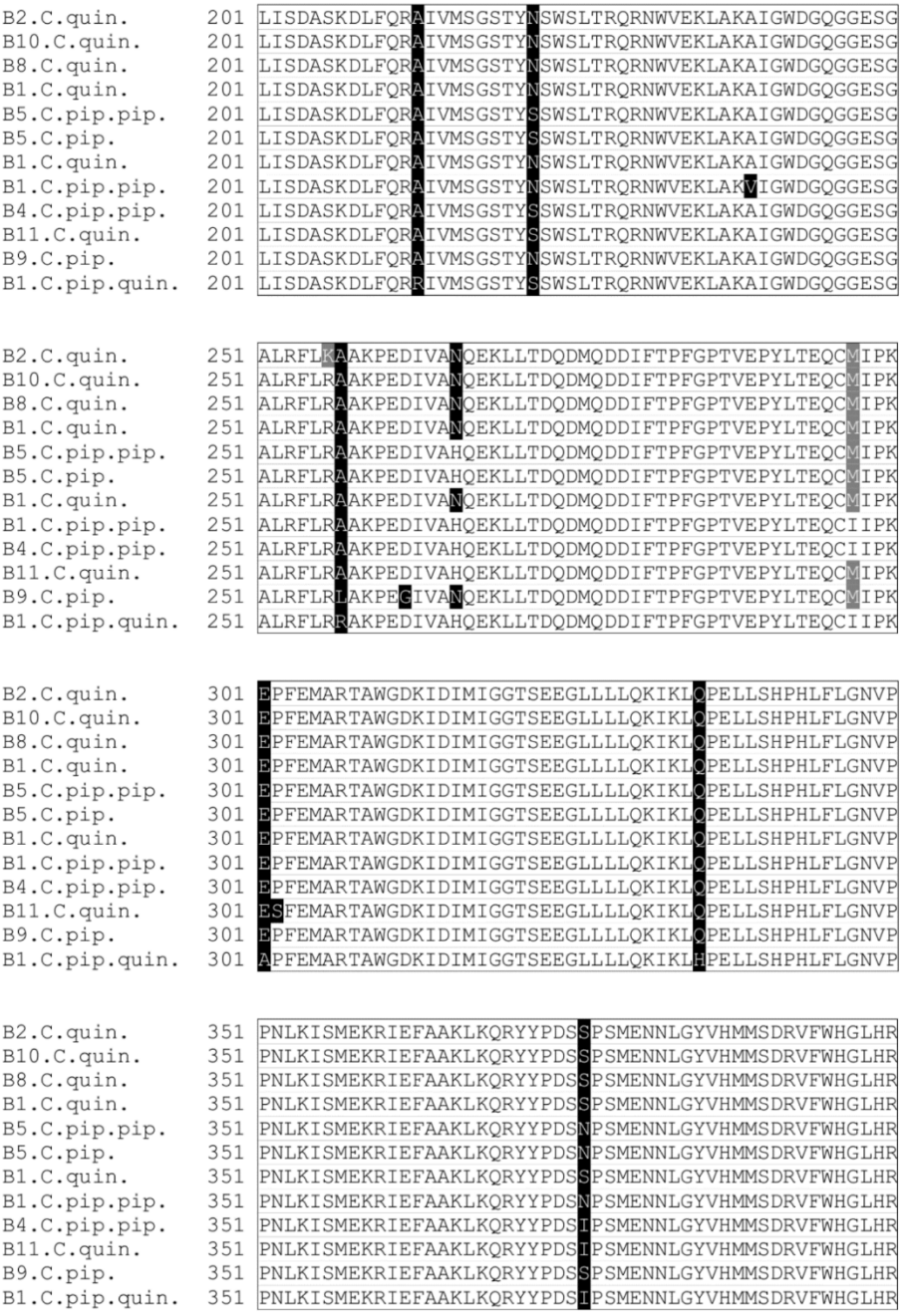
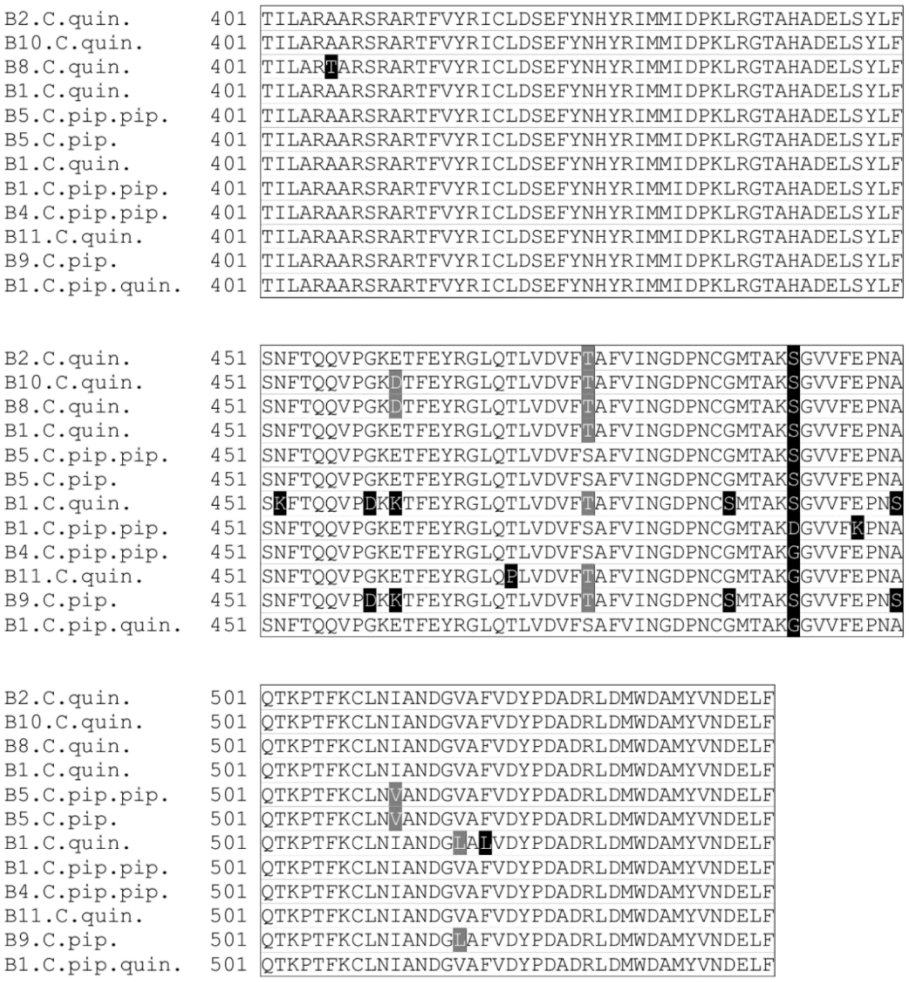
 Acetylcholine	 Juvenile Hormone II	 Example of a mid-chain food ester, Hexyl propanoate
 malonate ion	 Example of a FAME, Methyl decanoate	 Example of a medium-chain-length di-acylglycerol, 1, 3-diacylglycerol (C8 acyl chain length)
 4-nitrophenyl acetate	 4-nitrophenyl octanoate	 HEPES
 4-nitrophenyl butyrate	 Temephos	 Diethylumbelliferyl phosphate
 Propoxur	 Example of one of the stereoisomers of the cypermethrin analogues	 Paraoxon-ethyl

Figure S2. Multiple sequence alignment of the B2 isoforms from the *Culex* assemblage*




B2.C.quin. 201 LISDASKDLFQRAIVMSGSTYNSWSLTRQRNWVEKLAKAIGWDGQGGESG
B10.C.quin. 201 LISDASKDLFQRAIVMSGSTYNSWSLTRQRNWVEKLAKAIGWDGQGGESG
B8.C.quin. 201 LISDASKDLFQRAIVMSGSTYNSWSLTRQRNWVEKLAKAIGWDGQGGESG
B1.C.quin. 201 LISDASKDLFQRAIVMSGSTYNSWSLTRQRNWVEKLAKAIGWDGQGGESG
B5.C.pip.pip. 201 LISDASKDLFQRAIVMSGSTYSSWSLTRQRNWVEKLAKAIGWDGQGGESG
B5.C.pip. 201 LISDASKDLFQRAIVMSGSTYSSWSLTRQRNWVEKLAKAIGWDGQGGESG
B1.C.quin. 201 LISDASKDLFQRAIVMSGSTYNSWSLTRQRNWVEKLAKAIGWDGQGGESG
B1.C.pip.pip. 201 LISDASKDLFQRAIVMSGSTYNSWSLTRQRNWVEKLAKVIGWDGQGGESG
B4.C.pip.pip. 201 LISDASKDLFQRAIVMSGSTYSSWSLTRQRNWVEKLAKAIGWDGQGGESG
B11.C.quin. 201 LISDASKDLFQRAIVMSGSTYSSWSLTRQRNWVEKLAKAIGWDGQGGESG
B9.C.pip. 201 LISDASKDLFQRAIVMSGSTYNSWSLTRQRNWVEKLAKAIGWDGQGGESG
B1.C.pip.quin. 201 LISDASKDLFQRRIVMSGSTYSSWSLTRQRNWVEKLAKAIGWDGQGGESG
B2.C.quin. 251 ALRFLKAAKPEDIVANQEKLLTDQDMQDDIFTPFGPTVEPYLTEQCMIPK
B10.C.quin. 251 ALRFLRAAKPEDIVANQEKLLTDQDMQDDIFTPFGPTVEPYLTEQCMIPK
B8.C.quin. 251 ALRFLRAAKPEDIVANQEKLLTDQDMQDDIFTPFGPTVEPYLTEQCMIPK
B1.C.quin. 251 ALRFLRAAKPEDIVANQEKLLTDQDMQDDIFTPFGPTVEPYLTEQCMIPK
B5.C.pip.pip. 251 ALRFLRAAKPEDIVAHQEKLLTDQDMQDDIFTPFGPTVEPYLTEQCMIPK
B5.C.pip. 251 ALRFLRAAKPEDIVAHQEKLLTDQDMQDDIFTPFGPTVEPYLTEQCMIPK
B1.C.quin. 251 ALRFLRAAKPEDIVANQEKLLTDQDMQDDIFTPFGPTVEPYLTEQCMIPK
B1.C.pip.pip. 251 ALRFLRAAKPEDIVAHQEKLLTDQDMQDDIFTPFGPTVEPYLTEQCIIPK
B4.C.pip.pip. 251 ALRFLRAAKPEDIVAHQEKLLTDQDMQDDIFTPFGPTVEPYLTEQCIIPK
B11.C.quin. 251 ALRFLRAAKPEDIVAHQEKLLTDQDMQDDIFTPFGPTVEPYLTEQCMIPK
B9.C.pip. 251 ALRFLRLAKPEGIVANQEKLLTDQDMQDDIFTPFGPTVEPYLTEQCMIPK
B1.C.pip.quin. 251 ALRFLRRAKPEDIVAHQEKLLTDQDMQDDIFTPFGPTVEPYLTEQCIIPK
B2.C.quin. 301 EPFEMARTAWGDKIDIMIGGTSEEGLLLLQKIKLQPELLSHPHLFLGNVP
B10.C.quin. 301 EPFEMARTAWGDKIDIMIGGTSEEGLLLLQKIKLQPELLSHPHLFLGNVP
B8.C.quin. 301 EPFEMARTAWGDKIDIMIGGTSEEGLLLLQKIKLQPELLSHPHLFLGNVP
B1.C.quin. 301 EPFEMARTAWGDKIDIMIGGTSEEGLLLLQKIKLQPELLSHPHLFLGNVP
B5.C.pip.pip. 301 EPFEMARTAWGDKIDIMIGGTSEEGLLLLQKIKLQPELLSHPHLFLGNVP
B5.C.pip. 301 EPFEMARTAWGDKIDIMIGGTSEEGLLLLQKIKLQPELLSHPHLFLGNVP
B1.C.quin. 301 EPFEMARTAWGDKIDIMIGGTSEEGLLLLQKIKLQPELLSHPHLFLGNVP
B1.C.pip.pip. 301 EPFEMARTAWGDKIDIMIGGTSEEGLLLLQKIKLQPELLSHPHLFLGNVP
B4.C.pip.pip. 301 EPFEMARTAWGDKIDIMIGGTSEEGLLLLQKIKLQPELLSHPHLFLGNVP
B11.C.quin. 301 ESFEMARTAWGDKIDIMIGGTSEEGLLLLQKIKLQPELLSHPHLFLGNVP
B9.C.pip. 301 EPFEMARTAWGDKIDIMIGGTSEEGLLLLQKIKLQPELLSHPHLFLGNVP
B1.C.pip.quin. 301 APFEMARTAWGDKIDIMIGGTSEEGLLLLQKIKLHPELLSHPHLFLGNVP
B2.C.quin. 351 PNLKISMEKRIEFAAKLKQRYYPDSSPSMENNLGYVHMMSDRVFWHGLHR
B10.C.quin. 351 PNLKISMEKRIEFAAKLKQRYYPDSSPSMENNLGYVHMMSDRVFWHGLHR
B8.C.quin. 351 PNLKISMEKRIEFAAKLKQRYYPDSSPSMENNLGYVHMMSDRVFWHGLHR
B1.C.quin. 351 PNLKISMEKRIEFAAKLKQRYYPDSSPSMENNLGYVHMMSDRVFWHGLHR
B5.C.pip.pip. 351 PNLKISMEKRIEFAAKLKQRYYPDSNPSMENNLGYVHMMSDRVFWHGLHR
B5.C.pip. 351 PNLKISMEKRIEFAAKLKQRYYPDSNPSMENNLGYVHMMSDRVFWHGLHR
B1.C.quin. 351 PNLKISMEKRIEFAAKLKQRYYPDSSPSMENNLGYVHMMSDRVFWHGLHR
B1.C.pip.pip. 351 PNLKISMEKRIEFAAKLKQRYYPDSNPSMENNLGYVHMMSDRVFWHGLHR
B4.C.pip.pip. 351 PNLKISMEKRIEFAAKLKQRYYPDSIPSMENNLGYVHMMSDRVFWHGLHR
B11.C.quin. 351 PNLKISMEKRIEFAAKLKQRYYPDSIPSMENNLGYVHMMSDRVFWHGLHR
B9.C.pip. 351 PNLKISMEKRIEFAAKLKQRYYPDSSPSMENNLGYVHMMSDRVFWHGLHR
B1.C.pip.quin. 351 PNLKISMEKRIEFAAKLKQRYYPDSIPSMENNLGYVHMMSDRVFWHGLHR

S7



***list of Genbank accession numbers for each sequence included in the alignment, from top to bottom:** CAA83643.1, ABO85013.1, ABO85012.1, XP_001864331.1, AGM48552.1, AAS72505.1, CAA83644.1, AGM48548.1, AGM48550.1, ABO85015.1, AFJ93717.1, AAA28289.1.

Chapter 4. The Evolution of a Juvenile Hormone Esterase Duplication into an Odorant Degrading Enzyme in *Drosophila melanogaster*

4.1. Journal article overview

In the 1970s, a model for enzyme evolution was proposed, whereby ancestral enzymes were multifunctional generalists that evolved into more specialized enzymes through, gene duplication and sequence divergence (55, 225). While this model has provided a basic framework for the process of enzyme evolution, advances in technology have enhanced our understanding of this process and the complexities that are involved (56, 64, 67, 69–73). However, many studies rely on directed evolution and there are still relatively few, well studied examples of natural evolution, particularly in eukaryotes, where slower rates of evolution and gene loss complicate such studies (64, 226, 227). In particular, the question of how structural and regulatory changes are related to enzyme evolution remains unanswered (227–230). In this chapter, I present an example where a duplication of a vital hormone regulatory enzyme, JHE from *D. melanogaster*, has evolved to become an ODE involved in chemosensation.

JHEs were historically thought to be highly specific enzymes due to their critical role in degrading juvenile hormone (JH), which marks the transition from the last juvenile stage to either the adult stage in hemimetabolous or to the pupal stage in holometabolous insects (231). They have been identified in more than twenty species of insect from at least six insect orders (232–239). Generally, these JHEs had only been tested against a range of JH forms and not a wider range of substrates, thus little information was known about their promiscuity (239–242). In this chapter, I demonstrate that present day *D. melanogaster* JHE (DmJHE) has sufficient substrate promiscuity with mid-long chain food esters for a duplicate to evolve a general ODE function. All insect JHEs have previously been classified into two distinct clades; clade G for the lepidopteran-type JHEs and F for the dipteran-type JHEs (30, 35). I performed a more detailed phylogenetic analysis of insect hormone/semiochemical processing enzymes including JHEs, which demonstrates a more complex situation: the JHE lineage predates the hemi/holometabolan split and several instances where JHEs have diverged, likely through gene duplication, into β-esterases, such as ODEs, have occurred. I suggest that there are four groups of sequences that include JHEs, three of which contain other closely related β-esterases. Homology modelling of DmJHE allowed comparison with *Manduca sexta* JHE (MsJHE), which demonstrated

key structural conservation despite low sequence identity. This is consistent with previous comparisons of insect JHEs that revealed a number of highly conserved active site residues (232, 243, 244).

ODEs are critical enzymes in both communication and chemotaxis in insects (245). They are predominantly expressed in the antennae and act by terminating the signal from odorants, preventing their accumulation and subsequent sensory adaptation (246). This allows insects to rapidly respond to stimuli in their environment. While very few have been characterized, insect ODEs have been shown to function as either specific or general ODEs (245). The majority of identified ODEs are thought to be specific ODEs, due to high activity against certain pheromones (247–249). However, there are two examples of general ODEs, *D. melanogaster* Esterase 6 (DmEST6) and *Spodoptera exigua* SexiCXE4, which act upon food and plant odorant esters (24, 249–251). The duplication of DmJHE (DmJHEdup) described in this chapter is shown to also possess a broad substrate range against food and plant esters. DmJHEdup along with DmEST6 are the only ODEs to be identified in *D. melanogaster*. I show that DmEST6 and DmJHEdup have complementary substrate ranges: with DmEST6 having an optimal activity with short-mid chain esters; and DmJHEdup with mid-long chain esters. Homology modelling of DmJHEdup and comparison with DmEST6 demonstrated both similarities and differences with DmEST6 that explains both their similar roles and distinct kinetics.

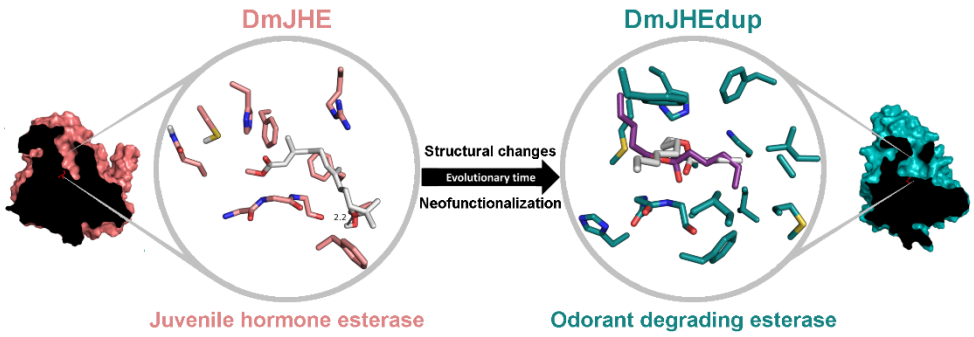
The evolution of DmJHEdup from an ancestral duplication of DmJHE is thus shown to represent a rare case where a specialist enzyme evolved into a general enzyme. While examples of this process are less common, it is not unprecedented (54, 252, 253). Biochemical comparisons demonstrated sufficient promiscuity in DmJHE and similarity with DmJHEdup to explain their relationship. Structural comparison of each enzymes homology models suggests a number of key changes, such as T309/L299, to the DmJHE active site that could have both abolished JHE activity and enhanced ODE activity resulting in DmJHEdup. While these structural changes to an ancestral JHE were likely early in the neofunctionalization process, the altered regulation and localization of DmJHEdup would have been a critical first step to provide immediate utility to such changes. This could have occurred via a transposon- mediated insertion

of an antennal-specific enhancer, eliminating the JHE expression profile and resulting in adult antennal expression (254–256).

4.2. Statement of contribution

The molecular basis for the neofunctionalization of the Juvenile Hormone Esterase Duplication in Drosophila

Davis H. Hopkins, Rahul V. Rane, Faisal Younus, Chris W. Coppin, Gunjan Pandey, Colin J. Jackson & John G. Oakeshott



This paper is under review in *Insect Biochemistry and Molecular Biology*. All enzyme assays, *in silico* docking, molecular comparisons, analysis of results, writing of the paper and discussion are my own work. Both Rahul Rane and John Oakeshott worked with me to select sequences for the phylogenetic analysis. While visualization and interpretation of the phylogenetic analysis was my own work, Rahul Rane conducted the computations to generate the phylogenetic tree. Faisal Younus was responsible for the expression of all enzymes used. Chris Coppin and Gunjan Pandey consulted on the appropriate methods for the experiments. John Oakeshott and Colin Jackson supervised the work, in addition to contributing to writing and editing the manuscript.

Relevant sections of the supporting information for this paper are presented following the manuscript.

The molecular basis for the neofunctionalization of the juvenile hormone esterase duplication in Drosophila

Davis H. Hopkins[a, b], Rahul V. Rane[b], Faisal Younus[a, b], Chris W. Coppin[b], Gunjan Pandey[b], Colin J. Jackson[a] & John G. Oakeshott[b]

[a]Research School of Chemistry, Australian National University, Canberra, Australian Capital Territory, 0200, Australia

[b]CSIRO Land and Water, Black Mountain, Canberra, Australian Capital Territory, 2601, Australia

Corresponding author: Davis H. Hopkins, u4843722@anu.edu.au, 0401038852

Abstract

The *Drosophila melanogaster* enzymes juvenile hormone esterase (DmJHE) and its duplicate, DmJHEdup, present ideal examples for studying the structural changes involved in the neofunctionalization of enzyme duplicates. DmJHE is a hormone esterase with precise regulation and highly specific activity for its substrate, juvenile hormone. DmJHEdup is an odorant degrading esterase (ODE) responsible for processing various kairomones in antennae. Our phylogenetic analysis shows that the JHE lineage predates the hemi/holometabolan split and that several duplications of JHEs have been templates for the evolution of secreted β-esterases such as ODEs through the course of insect evolution. Our biochemical comparisons further show that DmJHE has sufficient substrate promiscuity and activity against odorant esters for a duplicate to evolve a general ODE function against a range of mid-long chain food esters, as is shown in DmJHEdup. This substrate range complements that of the only other general ODE known in this species, Esterase 6. Homology models of DmJHE and DmJHEdup enabled comparisons between each enzyme and the known structures of a lepidopteran JHE and Esterase 6. Both JHEs showed very similar active sites despite low sequence identity (30%). Both ODEs differed drastically from the JHEs and each other, explaining their complementary substrate ranges. A particular amino acid change is identified which we propose was responsible for a critical early step in the neofunctionalization of DmJHEdup. Our results provide key insights into the process of neofunctionalization and the structural changes that can be involved.

Keywords

Neofunctionalization; Structural evolution; Juvenile hormone esterase; Odorant degrading enzyme.

Abbreviations[1]

1. Carboxylesterase – CBE, *Drosophila melanogaster* – Dm; *Manduca sexta* – Ms, Juvenile hormone esterase – JHE, Odorant degrading esterase – ODE, Esterase 6 – EST6

1. Introduction

Adaptation and the development of new physiological functions are driven by a combination of structural changes to proteins and changes in gene regulatory sequences (Carroll, 2005). There has been much discussion about how these two factors interact to effect the neofunctionalization of enzyme duplicates, particularly in higher eukaryotes, where tissue and temporal expression can be so varied (Carroll, 2005; Hoekstra and Coyne, 2007; Juneja et al., 2016; Tangwancharoen et al., 2018). The evolution of insect hormone and semiochemical processing enzymes provides an ideal system to study this interaction due to their reliance on both precise regulation and specific *in vivo* biochemistry (Oakeshott et al., 2005). Earlier work identified a group of four related clades of insect carboxylesterases (CBEs) that generally have roles in the processing of the sesquiterpenoid ester, juvenile hormone (JH), which plays a vital role in triggering metamorphosis, or various pheromone or food esters received by sensory tissues, which trigger various behavioral changes (Oakeshott et al., 2010; Sadd et al., 2015) . The temporal and tissue specificity required of the JH esterases (JHEs) is very different from that required of the odorant degrading esterases (ODEs) and their respective physiological roles impose very distinct requirements on their substrate specificities and kinetics (Oakeshott et al., 2005).

JHEs have been identified in more than twenty species of insect from at least six insect orders (El-Sheikh, 2015; Elayidam and Muraleedharen, 2008; Kamita et al., 2011; Kamita and Hammock, 2010; Kontogiannatos et al., 2013, 2011; Valaitis, 1991; Zhu et al., 2017), with the best characterized enzymes, from Lepidoptera and Diptera, falling into two different clades; clade G for the lepidopteran JHEs and F for the dipteran JHEs. These JHEs have been defined biochemically, in terms of their activity, by demonstrating a high k_{cat}/K_M against JH, predominantly due to a low apparent K_M, in the presence and absence of a carrier protein (Hammock, 1985). Notwithstanding their distribution across different phylogenetic clades, a range of active site residues are conserved amongst the JHEs, including RF, DQ, E, GxxHxxD/E, T and in particular a GQSAG which defines the residues around the catalytic serine (Kamita and Hammock, 2010; Takuya Tsubota et al., 2010). Intriguingly, the *Drosophila melanogaster* genome project discovered a closely related duplicate of its JHE gene,

(*DmJHE*), immediately upstream of it (Campbell et al., 2001). Subsequent physiological and behavioral work showed the product of the duplicate gene (DmJHEdup) functioned not as another JHE but as an antennal ODE active against certain food acetates, including isoamyl acetate, ethyl butyrate and ethyl propionate (Steiner et al., 2017). DmJHEdup does not possess the GQSAG motif and has been shown *in vitro* to have negligible activity with JHIII(Crone et al., 2007), however, the wrong start codon for DmJHEdup was used in the latter study, producing an unstable enzyme and compromising this result (Crone et al., 2007). Phylogenetic analysis of dipteran JHEs and JHEdups suggests the JHEdup duplication occurred early in the Brachycera (true fly) suborder, ie after the higher Diptera (flies) diverged from the lower Diptera (mosquitoes) (Steiner et al., 2017).

The only other well characterized *D. melanogaster* CBE in the four clades in question is Esterase 6 (DmEST6), which also acts as an antennal ODE (Younus et al., 2017, 2014), albeit it may also have other functions (Chertemps et al., 2012). DmEST6 lies in clade E, secreted β-esterases, which contains secreted CBEs from a range of insect orders that are involved in degrading hormones, pheromones and odorants (Oakeshott et al., 2010, 2005). Comprehensive biochemical analysis of DmEST6 shows it has relatively high activity against short-mid chain food esters and a crystal structure of the enzyme has provided key insights into the biophysical basis of its ODE function (Younus et al., 2017).

Here we present a revamped phylogenetic analysis of the JHE- and ODE-containing clades of insect CBEs, a detailed biochemical comparison of the activities of DmJHE and DmJHEdup with a range of esters, and a comparison of homology modelled structures for DmJHE and DmJHEdup and the known structures of a lepidopteran JHE and DmEST6. These analyses allow us to infer the key biochemical and structural differences among the enzymes and the sequence differences underpinning their functional divergence. Combining these results with recently published transcriptomic data for the enzymes allows us to infer key structural and regulatory changes associated with the neofunctionalization of DmJHEdup (Steiner et al., 2017; Younus et al., 2014).

2. Materials and Methods

2.1. Phylogenetic analysis of DmJHE and DmJHEdup

A phylogenetic tree was constructed using a total of 47 sequences obtained from official gene sets of the respective species on NCBI, two of which (AmelCCEFYa, AmelCCEFXa) were manually curated from GB11403 and GB18660 (BeeBase), respectively. *Drosophila melanogaster* acetylcholinesterase was used as an outgroup. Sequences include: all identified hormone/semiochemical CBEs from *Helicoverpa armigera, Bombyx mori* and *D. melanogaster*; homologous sequences to DmJHE and DmJHEdup in other insect orders; identified insect JHEs; and a selection of sequences that defined the clades D, E, F and G in previous phylogenies (Oakeshott et al., 2010, 2005). The protein alignments were carried out using the MAFFT program (Katoh et al., 2005) ("--localpair") and trimmed with trimAL (Capella-Gutierrez et al., 2009) to remove gaps. The phylogeny was constructed with IQ-TREE (Nguyen et al., 2015) using the "LG+R5" model identified to be the best-fit model by ModelFinder (Kalyaanamoorthy et al., 2017). The phylogenies were then plotted using the iTOL portal (Letunic and Bork, 2016).

2.2. Activity Assays

The expression of DmJHE and DmJHEdup (FlyBase Release 6 version) using the baculovirus system has been previously described (Younus et al., 2014). Enzyme concentration was determined through titration using varying concentrations of the inhibitor, Dibrom (Sigma-Aldrich) and by incubating at room temperature for 20 minutes before using α-napthyl acetate (Sigma-Aldrich) to determine the remaining activity. All assays were performed at room temperature.

2.2.1. Artificial esters

Michaelis-Menten kinetics were determined for both DmJHE and DmJHEdup with 4-nitrophenyl esters with saturated acyl chain lengths varying from 2 – 12 carbon atoms long (Sigma-Aldrich) as well as with α-naphthyl acetate. Assays were performed in triplicate with varying substrate concentrations (0 – 2 mM) in 25 mM Tris-HCl buffer (pH 7.5) with ethanol (5% v/v). Enzyme concentrations varied from 10 – 25 nM. The

formation of 4-nitrophenol was measured at 405 nm and the concentration was determined using a molar extinction coefficient (ε = 18400 M^{-1} cm^{-1}). The formation of α-naphthol was measured at 325 nm and the concentration was determined using a molar extinction coefficient (ε = 1114 M^{-1} cm^{-1}). Michaelis-Menten kinetics were calculated from the initial velocity data using nonlinear regression.

2.2.2. Odorant and hormone esters

The activities of both DmJHE and DmJHEdup were determined against a range of odorant and hormone esters (Sigma-Aldrich) using gas chromatography-mass spectrometry (GC-MS) assays and GC with flame ionization detection (GC-FID) assays modified from Younus et al. (2014) (Younus et al., 2014). Briefly, substrate concentration was kept at 200 µM in 25 mM Tris-HCl buffer (pH 7.5) with ethanol (5% v/v). Enzyme concentrations were varied from 0.1 - 90 nM. Reactions were stopped at specific time points through addition of ice-cold hexane (containing 200 µM heptanone as an external standard) and vortexed for 10 minutes. The hexane layer was extracted and substrate loss was determined using GC-MS and GC-FID.

2.2.3. Comparative chiral activity assay

Qualitative JHE preference for Juvenile hormone III isomer was determined through a high-performance liquid chromatography (HPLC) chiral separation technique modified from Ichikawa et al. (2007) (Ichikawa et al., 2007). Briefly, activity assays were performed as described above. Extracted samples were separated using a chiralpak IA column (DAICEL Co., Ltd) under normal-phase conditions; mobile phase: hexane/EtOH = 99.5/0.5; flow rate 0.5 ml/min; detection: UV 230 nm. Isomers were identified through HPLC profile comparison with Ichikawa et al. (2007) (Ichikawa et al., 2007).

2.2.4. Inhibition assays

The inhibition constant, K_i, of DmJHE and DmJHEdup towards the odorant and hormone esters was determined using a competitive assay modified from Younus et al. (2014) and based on the concept that competitive substrates can be treated as inhibitors of each other (Cornish-Bowden, 1995; Younus et al., 2014). Briefly, the full kinetics towards the substrate α-naphthyl acetate were determined for both DmJHE and DmJHEdup in 25 mM Tris-HCl buffer (pH 7.5) with ethanol (5% v/v). Then,

competitive assays were performed between each of the odorant and hormone esters, with a range of concentrations (0 – 32 mM) against a single concentration of α-napthyl acetate (500μM). K_i was determined using the following formula:

$$K_{\mathrm{i}} = \frac{\frac{[\mathrm{I}]}{(1-\alpha)} - [\mathrm{I}]}{1 + \frac{[\mathrm{S}]}{K_{\mathrm{M}}}}$$

Where $\alpha = v_i / v_0$ = relative activity; v_i = the initial velocity at a given substrate concentration, [S], and in the presence of inhibitor at a certain concentration, [I]; and v_0 = the initial velocity at the same [S] in the absence of inhibitor (Cornish-Bowden, 1995; Segel, 1993).

2.3. Homology modelling and docking

The structural modelling of both DmJHE and DmJHEdup was performed using the ROBETTA server (Kim et al., 2004). These were compared with other homology models using the QMEAN server and found to be of the highest quality (Benkert et al., 2008). These structures were compared with structurally determined enzymes with similar functions using PyMol, which was also used to make all protein structure images (Schrodinger LLC, 2010). To determine the potential acyl-enzyme complexes of DmJHEdup and DmJHE formed with JHIII, automated covalent docking using flexible residues was performed using AutoDock4 as described by Bianco et al. (2016) (Bianco et al., 2016; Morris et al., 2009). To determine the potential tetrahedral intermediates formed between a selection of odorant esters and DmJHE and DmJHEdup, covalent docking was conducted with CovDock (Zhu et al., 2014).

3. Results

3.1. Evolutionary relationships of DmJHE and DmJHEdup

A phylogeny was constructed using 47 functionally annotated sequences of insect CBEs which preliminary analyses indicated would sit in one of the four clades originally identified as containing hormone and semiochemical processing enzymes (**Figure 1**). Those four original clades were: clade D, integument esterases; clade E, secreted β-esterases; clade F, dipteran-type JHEs; and clade G, lepidopteran-type JHEs (Oakeshott et al., 2010, 2005). The 47 sequences were drawn from seven insect orders, namely the Diptera, Coleoptera, Lepidoptera, Hemiptera, Hymenoptera, Orthoptera and Blattodea. Each major branch in our new phylogeny is supported by a bootstrap value of at least 60, with the majority of branches supported by bootstrap values above 98. Our phylogeny breaks the CBEs analyzed into seven distinct groups. The integrity of clades G and D are retained and clade G is confirmed as the ancestral clade. However, clades E and F containing DmJHE, DmJHEdup and DmEST6 are now re-arranged and partitioned into five subclades: E-1; E-2; F-1; F-2; and F-3.

There are now four robust groups of sequences that include JHEs: the first, clade G, contains lepidopteran JHEs and β-esterases; the second, clade F-1, contains solely coleopteran JHEs; the third, clade F-2, contains DmJHE and DmJHEdup, other dipteran and orthopteran JHEs and other dipteran β-esterases; and the fourth, clade F-3, contains hemipteran, hymenopteran and blattodean JHEs along with other dipteran, lepidopteran and hymenopteran β-esterases. The branching of these groups and their divergence order is not related to the splits of the organismal orders, with two groups (F-2 and F-3) containing JHEs from both hemi- and holometabolous insects (Misof et al., 2014). This establishes that the JHE lineage predates the hemi/holometabolan split and has given rise to many different subclades with JHE functions (Misof et al., 2014).

Although the integrity of clade D, integument esterases, is retained, its evolutionary relationship with other clades is altered in this phylogeny (Oakeshott et al., 2010, 2005). Before clade D was thought to exist as a monophyletic clade distinct from

clades E and F, whereas this phylogeny suggests that it evolved within the JHE lineage.

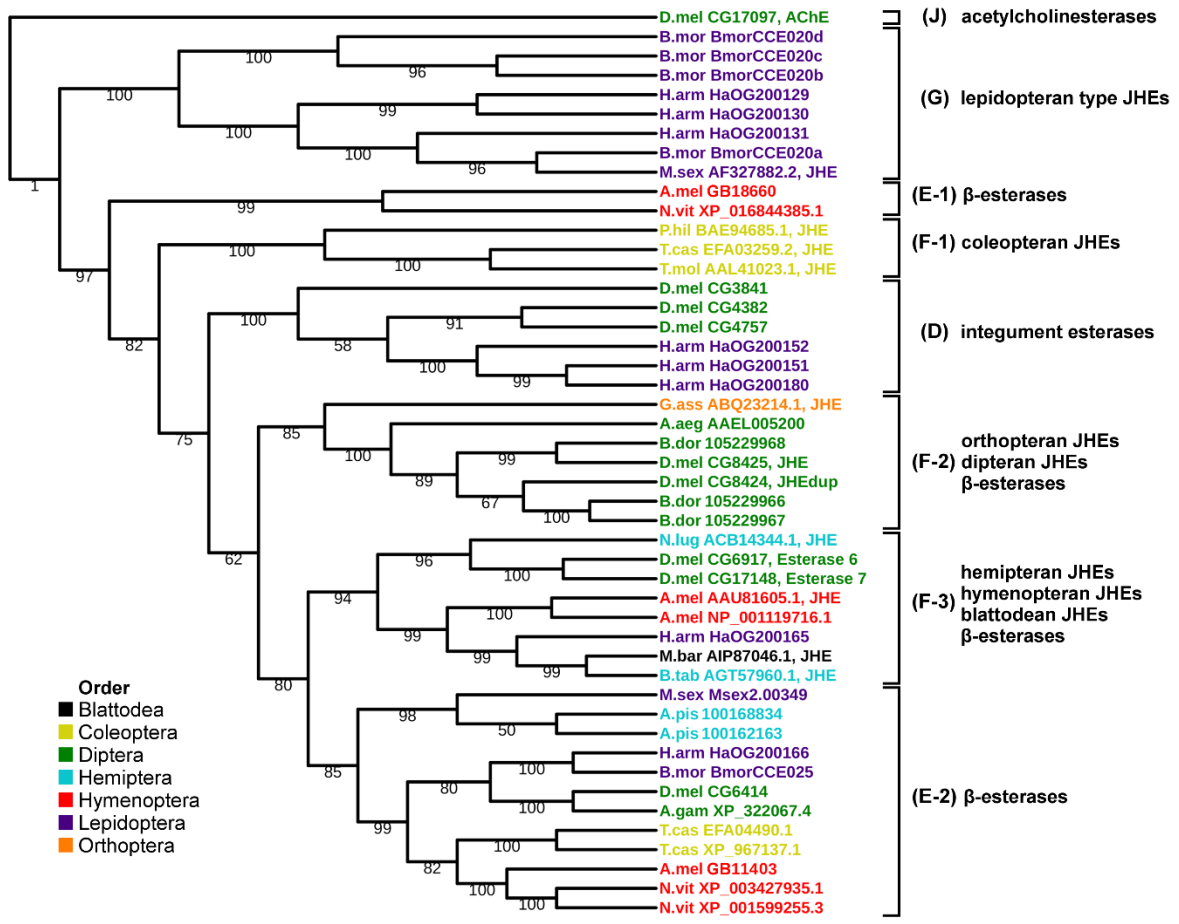


Figure 1: A phylogenetic tree of the evolutionary relationship between insect JHEs and other hormone/semoichemical processing CBEs. *Drosophila melanogaster* acetylcholinesterase was used as an outgroup. Sequences include all identified hormone/semiochemical CBEs from *Helicoverpa armigera, Bombyx mori* and *D. melanogaster*, homologous sequences to DmJHE and DmJHEdup in other insect orders, identified insect JHEs and CBEs and a selection of sequences that defined the clades D, E, F and G in previous phylogenies (Oakeshott et al., 2010, 2005). Sequences are first labelled by a shorthand for species name: D.mel, *D. melanogaster*; B.mor, *B. mori*; H.arm, *H. armigera*; M.sex, *Manduca sexta*; A.mel, *Apis mellifera*; N.vit, *Nasonia vitripennis*; P.hil, *Psacothea hilaris*; T.cas, *Tribolium castaneum*; T.mol, *Tenebrio molitor*; G.ass, *Gryllus assimilis*; A.aeg, *Aedes aegypti*; B.dor, *Bactrocera dorsalis*; N.lug, *Nilaparvata lugens*; M.bar, *Macrotermes barneyi*; B.tab, *Bemisia tabaci*; A.pis, *Acyrthosiphon pisum*; and A.gam, *Anopheles gambiae*. Sequences are then labelled either by the appropriate genome annotation number, by genbank/NCBI gene ID, or NCBI reference number, followed by generic name, if appropriate. Sequences are colored based on insect order: Diptera, green; Lepidoptera, purple; Hymenoptera, red; Coleoptera, yellow; Orthoptera, orange; Hemiptera, blue; and Blattodea, black. Sequences are grouped based on suggested phylogenetic clades shown in brackets. Percentage bootstrap values are shown prior to each node. Phylogenetic groups are labelled.

While the previous phylogenetic studies suggested clade E is the only secreted β-esterase clade, our phylogenetic analysis suggests a more complex situation. There are two groups, E-1 and E-2, that contain solely secreted β-esterases and three groups, G, F-2 and F-3, with both β-esterases and JHEs. While DmJHEdup and

DmEST6 share a common function as ODEs in *D. melanogaster,* they occur in distinct groups (F-2 and F-3, respectively), both of which contain insect JHEs. This demonstrates that JHEs and JHE-like CBEs can act as templates for the development of enzymes with secreted β-esterase functions over evolutionary time.

3.2. Kinetic comparison between DmJHEdup and DmJHE

We assayed DmJHEdup and DmJHE against a range of artificial 4-nitrophenyl esters with varying acyl chain lengths at a range of substrate concentrations, allowing the determination of full Michaelis-Menten kinetics (**Table 1**). DmJHEdup has an optimal k_{cat} at an acyl chain length of six, where it is >15-fold higher than that of DmJHE. DmJHE has an optimal k_{cat} value with an acyl chain length of only two, albeit DmJHEdup still has similar activity with that substrate. Both enzymes have their lowest K_M values, in the low micromolar to high nanomolar range, somewhat lower for DmJHEdup than DmJHE, with longer (8 or 10) acyl moiety substrates. Specificity constants are also highest for both enzymes on these substrates, those for DmJHEdup nearly ten-fold higher than those for DmJHE.

Table 1. Kinetic parameters of DmJHEdup and DmJHE with 4-nitrophenyl esters with varying acyl chain length.

4-nitrophenyl acyl chain length	k_{cat} (min^{-1})		K_M (μM)		k_{cat}/K_M ($M^{-1}.s^{-1}$)	
	DmJHEdup	DmJHE	DmJHEdup	DmJHE	DmJHEdup	DmJHE
2	834 ± 38	816 ± 59	$2.14\times10^{2} \pm 0.14\times10^{2}$	$2.22\times10^{3} \pm 0.20\times10^{3}$	$6.50\times10^{4} \pm 0.14\times10^{4}$	$6.13\times10^{3} \pm 0.09\times10^{3}$
4	1190 ± 20	123 ± 0.4	$1.36\times10^{1} \pm 0.03\times10^{1}$	$1.84\times10^{2} \pm 0.01\times10^{2}$	$1.46\times10^{6} \pm 0.02\times10^{6}$	$1.12\times10^{4} \pm 0.01\times10^{4}$
6	1540 ± 16	90.4 ± 3.0	$8.10\times10^{0} \pm 0.02\times10^{0}$	$3.87\times10^{1} \pm 0.36\times10^{1}$	$3.16\times10^{6} \pm 0.03\times10^{6}$	$3.91\times10^{4} \pm 0.25\times10^{4}$
8	554 ± 11	182 ± 2	$1.38\times10^{0} \pm 0.08\times10^{0}$	$3.03\times10^{0} \pm 0.11\times10^{0}$	$6.73\times10^{6} \pm 0.50\times10^{6}$	$9.98\times10^{5} \pm 0.27\times10^{5}$
10	268 ± 2	345 ± 7	$5.40\times10^{-1} \pm 0.92\times10^{-1}$	$6.55\times10^{0} \pm 0.40\times10^{0}$	$8.41\times10^{6} \pm 1.36\times10^{6}$	$8.79\times10^{5} \pm 0.42\times10^{5}$
12	79.1 ± 0.5	101 ± 2	$1.44\times10^{0} \pm 0.13\times10^{0}$	$7.43\times10^{0} \pm 1.12\times10^{0}$	$9.20\times10^{5} \pm 0.89\times10^{5}$	$2.29\times10^{5} \pm 0.30\times10^{5}$

Assays were also carried out against a racemic mixture of JHIII (the major form of JH for most insects, including *Drosophila* (Kamita et al., 2003; Kamita and Hammock, 2010; Noriega, 2014)) and some other natural bioactive esters, mainly food odorants to which *D. melanogaster* is known to respond (Younus et al., 2017). These assays were restricted by substrate solubility, so activities were determined at a set substrate concentration (200 μM) (**Table 2**). To determine the Michaelis-Menten constant, K_M, we utilized the fact that in a reaction including two CBE substrates, each can be treated as a competitive inhibitor of the other and that the determined K_i for a substrate is

equivalent to its K_M (Cornish-Bowden, 1995; Younus et al., 2014). Thus, while we were unable to determine Michaelis-Menten kinetics directly, we were able to determine the K_i with JHIII and the other bioactive esters by utilizing a competitive inhibition assay between each substrate and the fluorometric, artificial substrate, α-naphthyl acetate (**Table 2**).

Table 2. The specific activities and inhibition constants determined for DmJHE and DmJHEdup against a range of food odorant, green plant odorant, pheromonal and hormonal esters. Substrates have been grouped based on both source and alcohol and acyl chain length (C_x-C_y, C_a-C_b, respectively).

		Specific activity (min^{-1})		K_i (μM)	
Source	**Substrates**	**JHEdup**	**DmJHE**	**JHEdup**	**DmJHE**
Food odorant	**Ethyl butyrate**	8.27 ± 4.29	4.92 ± 0.30	62600 ± 5600	11300 ± 200
C_1-C_3, C_1-C_4	**Propyl propionate**	162 ± 44	310 ± 34	395000 ± 8000	18300 ± 800
Food odorant	**Pentyl hexanoate**	2050 ± 690	0	588 ± 78	36.6 ± 3.5
C_5-C_8, C_4-C_6	**Octyl butyrate**	2540 ± 490	186 ± 18	257 ± 15	330 ± 27
Food odorant	**Pentyl formate**	240 ± 55	13.6 ± 3.6	20600 ± 2400	9470 ± 360
C_5-C_{10}, C_1-C_2	**Decyl acetate**	316 ± 66	0	272 ± 52	55.1 ± 10.0
Plant odorant	**Methyl decanoate**	42.0 ± 12.9	492 ± 16	1990 ± 360	46.1 ± 9.1
C_1-C_4, C_{10}-C_{12}	**Methyl myristate**	0	12.2 ± 1.2	0	13.9 ± 0.6
Food odorant aromatic ester	**Phenethyl acetate**	107 ± 35	38.9 ± 24.4	3370 ± 320	10400 ± 1200
Insecticide	**Methoprene**	0	0	226 ± 78	203 ± 69
Hormone	**Methyl farnesoate**	0	6.64 ± 2.70	93.2 ± 12.9	5.88 ± 0.20
	JHIII (racemate)	0	87.7 ± 4.5	0	3.35 ± 0.09
Pheromone	***cis*-Vaccenyl acetate**	0	0	0	0

Consistent with its known physiological function, DmJHE has significant activity and a low micromolar K_i with JHIII (**Table 2, Figure 2**). By contrast, DmJHEdup has no measurable activity or K_i with this substrate, confirming that it has no role as a JHE physiologically. Chiral separation of the reaction between DmJHE and racemic JHIII confirmed that DmJHE preferentially reacts with the biologically active (10R)-2E JHIII enantiomer (**Figure S1.**) (Campbell et al., 1998; Crone et al., 2007). Notably, DmJHE shows less or no activity with the JH precursor, methyl farnesoate, and the JH analog, methoprene, although it still has K_i values in the micromolar range with both these substrates (**Table 2, Figure 2**). The K_i result with methoprene is noteworthy in light of its known role as an insect growth regulator insecticide with an antagonist mode of action (Barry et al., 2008; Wilson and Fabian, 1986; Yin et al., 1987). DmJHEdup has no measurable activity with either methyl farnesoate or methoprene and K_i values in the high micromolar range with both, again consistent with it having a completely different physiological function from DmJHE (**Table 2, Figure 2**).

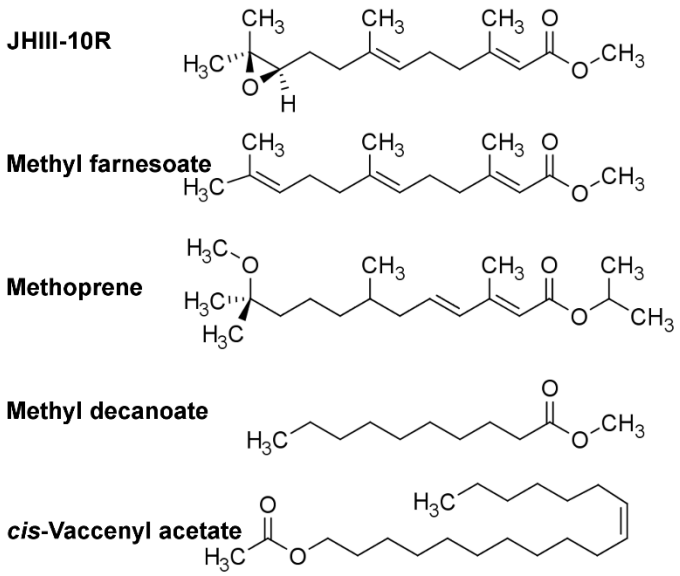


Figure 2. The chemical structures of a number of key substrates reacted with DmJHE and DmJHEdup.

The activities of both enzymes with the other bioactive esters confirm their respective structural substrate preferences. Of these esters, DmJHE shows the highest activity with methyl decanoate, an ester of a similar length and configuration as JHIII but measurable activity with some other esters with shorter acyl chains (**Table 2, Figure 2**). By contrast, DmJHEdup shows minimal activity with methyl decanoate, relatively low activity towards esters with shorter acyl and alcohol groups such as ethyl butyrate and propyl propionate, but high activity towards esters with mid-long length acyl and alcohol groups such as pentyl hexanoate and octyl butyrate (**Table 2**). Again, consistent with the 4-nitrophenol ester results, both enzymes give high micromolar/low millimolar K_M values with most of the esters with either short or mid-length acyl and alcohol groups. Neither enzyme displayed any activity against the important sex pheromone, *cis*-vaccenyl acetate (**Table 2, Figure 2**) (Chertemps et al., 2012; Younus et al., 2017).

3.3. Structural comparison between DmJHEdup, DmJHE and other JHEs

Homology models were generated for both DmJHEdup and DmJHE using the ROBETTA server and α-Esterase 7 from *Lucilia cuprina* as the template (PDB: 5CH3 and PDB: 5TYM, respectively) (Correy et al., 2016; Kim et al., 2004). The quality of each model was assessed using the QMEAN server (Benkert et al., 2008). The QMEAN6 Z-scores of DmJHEdup and DmJHE were -1.80 and -1.29, respectively, indicating models of sufficient quality to interpret the kinetic and phylogenetic results in a structural sense and draw comparisons with known insect CBE structures such as MsJHE and DmEST6 (Wogulis et al., 2006; Younus et al., 2017).

Both the modelled structure of DmJHE and the crystal structure of MsJHE have binding pockets (Wogulis et al., 2006) in the form of long, narrow tunnels in the same orientation (**Figure 3A**). Covalent docking simulations (**Figure 3B, D**) show both also utilize similar residues to interact with JHIII. These include two phenylalanine residues (F251 and F427 in DmJHE and F259 and F425 in MsJHE) that form π-π interactions with the conjugated portions of JHIII and a threonine residue (T309 in DmJHE and T314 in MsJHE) that forms hydrogen bonding interactions with the epoxide moiety of JHIII. These residues are conserved among other insect JHEs and are known to be key in JHIII binding (Kamita et al., 2010, 2003; Kamita and Hammock, 2010). The glutamine residue (Q217 in DmJHE and Q225 in MsJHE) that precedes the catalytic serine in the primary sequence is also highly conserved between insect JHEs (Kamita et al., 2010, 2003; Kamita and Hammock, 2010). In MsJHE, Q225 is shown to be in hydrogen bonding distance to the bound JHIII's acyl group (**Figure 3D**); this interaction may be important in orienting and stabilizing JHIII. In DmJHE, the docking simulation did not orient the conserved glutamine to interact with JHIII, however none of the surrounding residues prevent interaction, suggesting this poor orientation may be an artifact from the simulation (**Figure 3B**). Thus, overall, both the position and nature of the key residues in the active site are consistent between the two enzymes even though they possess very low sequence identity (30%) (Altschul et al., 1990).

While the modelled structures of DmJHEdup and DmJHE use the same template, their structures show significant differences. In contrast to the long, narrow binding pocket of DmJHE, the binding pocket of DmJHEdup is wider and more open, consistent with

the broader substrate specificity observed in the kinetic data (**Table 1, Table 2, Figure 3A**). The CASTp server determined that the volumes of the binding pockets of DmJHEdup and DmJHE are 1374.1 $Å^3$ and 460.0 $Å^3$, respectively (Dundas et al., 2006). Superimposition of JHIII over the active site of DmJHE also shows why the latter does not interact with JHIII. (**Figure 3C**). All the key residues for JHIII binding in DmJHE are either mutated or shifted in DmJHEdup: F251 and Q217 in DmJHE are replaced by A240 and H206 in DmJHEdup, respectively; L299 in DmJHEdup replaces T309 in DmJHE and directly clashes with JHIII binding; and while F427 in DmJHE aids JHIII binding, the equivalent residue in DmJHEdup, F414, is shifted away. Both DmJHE and DmJHEdup possess predominantly hydrophobic binding pockets, however all of the residues of DmJHEdup are shifted further from the catalytic serine, which would allow easier diffusion of substrates into the active site and accommodate the broader substrate range observed for DmJHEdup.

3.4. Biochemical and structural comparison between DmJHEdup and DmEST6

We also compared the biochemistry of DmJHEdup with that of the only other well studied *Drosophila* ODE, DmEST6 (Younus et al., 2017). Both enzymes show poor activity with odorant esters that possess small alcohol groups, irrespective of the length of the acyl group (Younus et al., 2017). However, DmEST6 has its greatest activity with esters with propionate acyl groups and alcohol groups from 3 to 8 carbon atoms long (Younus et al., 2017), whereas DmJHEdup has greatest activity against substrates with longer acyl groups, butyrate and hexanoate, albeit with a similar preference for alcohol groups (**Table 2**). The differences in acyl group preferences suggest the two enzymes may have complementary roles in processing the array of food esters to which *D. melanogaster* responds.

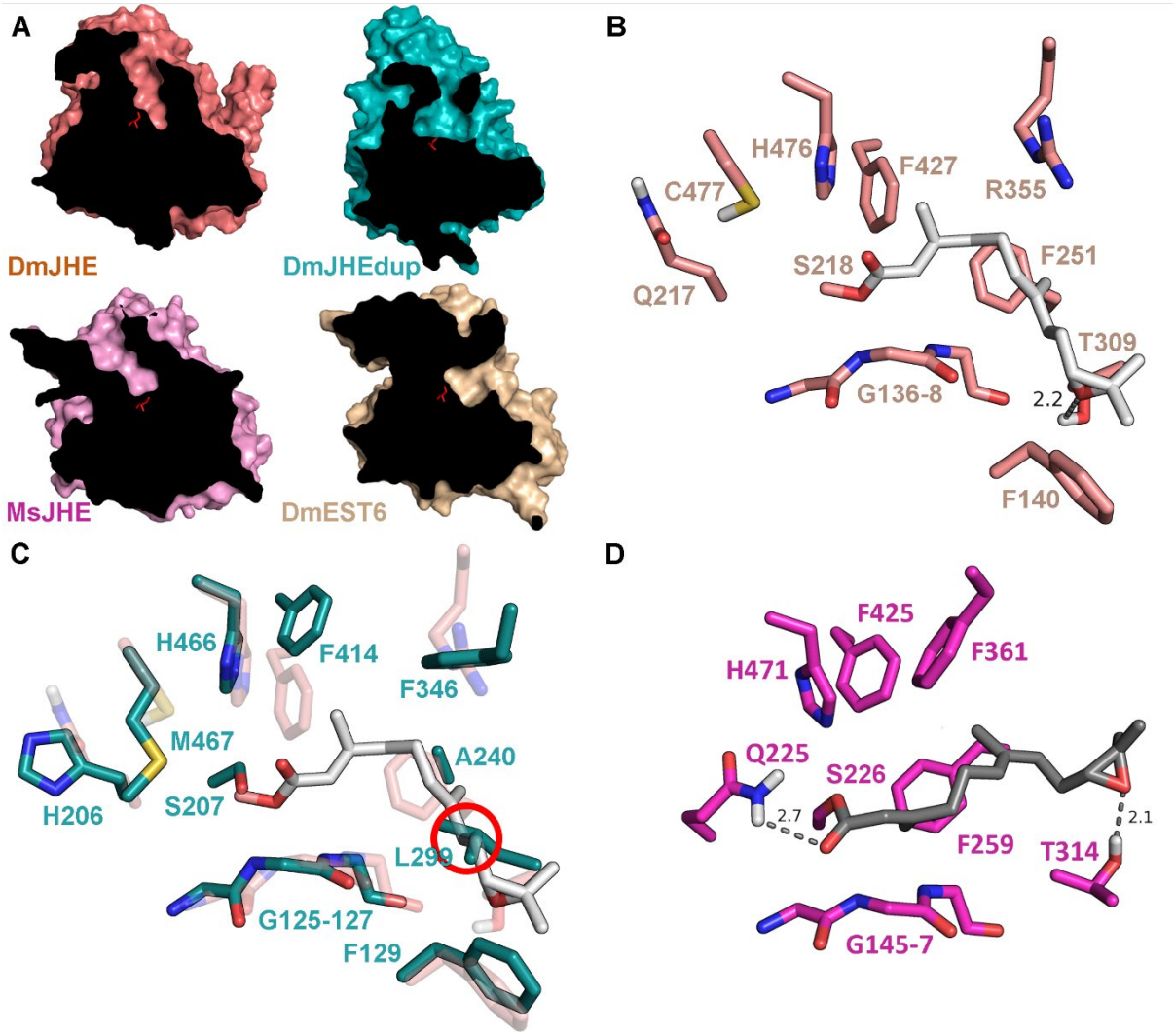


Figure 3. (A) A comparison between the predicted binding pockets of DmJHE (red) and DmJHEdup (cyan) against the pockets of the solved structures of MsJHE (pink) and DmEST6 (sand). **(B)** The docked interaction between JHIII and the modelled active site of DmJHE indicating key residues involved in binding. **(C)** The active site of the modelled DmJHEdup indicating key residues. A transparent representation of the DmJHE active site with JH bound is overlaid and the location of a key clash in the DmJHEdup structure is indicated. **(D)** The covalently docked interaction between the crystal structure of MsJHE and JHIII with key residues involved in binding labelled.

The modelled binding pocket of DmJHEdup was also compared to the crystal structure of DmEST6. Both enzymes share a different active site entrance from the JHEs (**Figure 3**). However, the two enzymes also differ significantly from each other in their binding pockets. The DmEST6 binding pocket is much smaller (227.8 Å^3) than that of DmJHEdup (1374.1 Å^3) and it also differs significantly in shape. The acyl binding pocket of DmEST6 consists of a range of bulky, hydrophobic residues such as F276 and F397 and is primarily restricted by W221 (**Figure 4B**), whereas the buried acyl

binding pocket of DmJHEdup is larger, being defined by smaller and more flexible hydrophobic residues such as I211, L238, A240, M241, and L299 (**Figure 4A**) and with no residue equivalent to W221 to restrict it. These differences could explain why DmJHEdup is better suited to substrates with longer acyl groups. The alcohol binding pockets of both enzymes include their active site entrances and are quite broad with both consisting of predominantly bulky, hydrophobic residues (**Figure 4A, B**). These similarities can explain their similar preferences for ester alcohol groups.

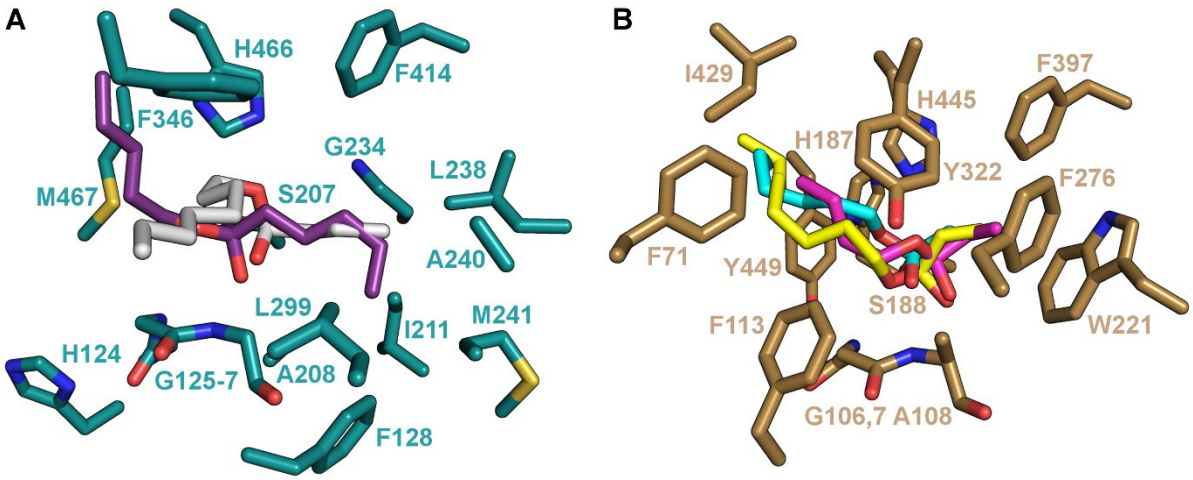


Figure 4. (A) A covalent docking simulation showing the tetrahedral intermediates formed between DmJHEdup and the efficiently hydrolyzed substrates octyl butyrate (grey) and pentyl hexanoate (purple). **(B)** A covalent docking simulation showing the tetrahedral intermediates formed between DmEST6 and the efficiently hydrolyzed substrates octyl propionate (yellow), butyl propionate (pink) and *trans*-2-hexenyl acetate (blue). Both simulations were aligned and have been shown from the same direction with the acyl binding pocket to the right and the alcohol binding pocket/active site entrance to the left.

4. Discussion

JH plays an essential role in the regulation of development in every insect which has been studied, covering a range of hemi- and holometabolous orders (Bai et al., 2007; Cornette et al., 2008; Hinton and Hammock, 2003; Kamita and Hammock, 2010; Li, 2007; Robinson et al., 1991; Teal et al., 2014). It also carries out similar functions in various other arthropods (Cusson et al., 1991; Jindra et al., 2013; Smykal et al., 2014). It is therefore assumed that JHE has also provided an essential function since early arthropod evolution (Cusson et al., 1991; Jindra et al., 2013; Smykal et al., 2014). While earlier phylogenetic studies, using fewer identified JHEs, suggested that they could be classified into at least two distinct CBE clades, our phylogeny now shows them to be spread across four CBE lineages, most of which include enzymes with non-JHE functions (Oakeshott et al., 2010, 2005). Moreover, the topology of these lineages does not correspond well with the topology of a phylogeny for the respective insect orders (Misof et al., 2014). We conclude that several different daughter lineages of the ancestral JHE have retained the JHE function, whilst also independently duplicating to evolve various other functions.

Significantly there is no evidence from any study for more than one functional JHE in a given insect, so we further suggest that there has been selection against the retention of JHE function in any duplicate of a functional JHE through the course of insect evolution (Crone et al., 2007; Gilbert et al., n.d.; Kamita and Hammock, 2010; Oakeshott et al., 2005; T. Tsubota et al., 2010; Tsubota and Shiotsuki, 2010). Further evidence for this is that transgenic manipulation of JHE activity or expression profiles has been shown to disrupt insect development (Hajós et al., 1999; Hammock et al., 1990).

DmJHEdup is a relatively recent example of the divergent evolution of a duplicate JHE, having arisen in the Brachycera lineage of the Diptera, before the emergence of the Schizophora (Steiner et al., 2017; Yeates and Wiegmann, 1999). In particular, we extend earlier work showing its function does not overlap with DmJHE. The previous work had shown that it has evolved a very different tissue and temporal expression profile from DmJHE**:** DmJHEdup is the most highly expressed CBE in the antenna

where DmJHE has low expression; and DmJHE has high expression in the adipose tissue during the pupal stage of development where DmJHEdup has no detectable expression (Younus et al., 2014). Our study shows that despite a relatively wide substrate range, at least *in vitro*, DmJHEdup also lacks detectable amounts of JHIII hydrolytic activity or even binding. Thus, DmJHEdup is incapable of interfering with the physiological function of DmJHE on both regulatory and structural grounds. Equally, we now know, from the work of Younus et al (2014), that DmJHEdup has the expression profile and, from the current study, that it has the kinetic capability to function as an ODE in a complementary fashion to the only other known Drosophila ODE, DmEST6 (Younus et al., 2017, 2014). DmJHEdup thus exemplifies the 'neofunctionalization' that is central in the theory of the evolution of new biochemical functions by gene duplication (Hahn et al., 2007; Kondrashov, 2012; Oakeshott et al., 1993).

The question then arises as to what sort of mutation could have occurred early in the life of the DmJHEdup gene that would have obliterated any physiologically relevant JHIII hydrolytic function whilst at the same time retaining some expression and activity as a template from which selection could then act to evolve its eventual role as an effective ODE? Without some such utility the cognate gene would accumulate disabling mutations through neutral evolution, becoming irreversibly 'pseudogenized' and eventually lost (Copley, 2010; DePristo et al., 2005; Li et al., 1981; Podlaha and Zhang, 2010).

We suspect that a crucial early step in the evolution of DmJHE would have been a cis-inherited regulatory change that eliminated the expression profile associated with *in vivo* JHIII degradation whilst bestowing the adult antennal expression from which the new ODE function might evolve. For example, an antennal-specific enhancer might have been inserted into the promoter of the gene (perhaps via a transposable element) in such a way as to disrupt an element required for the ancestral JHE expression profile (Kidwell and Lisch, 1997; Levin and Moran, 2011; Wittkopp and Kalay, 2012). There are several precedents for transposon mediated insertions into eukaryote promoters that eliminate previous aspects of expression and create qualitatively new ones (Chung et al., 2007; Jeong et al., 2006; Li et al., 2007; Schlenke and Begun, 2004). The promiscuous activities of the duplicated JHE for a range of mid-long chain

esters might then have had some immediate ODE utility and at the same time served as a template for further evolution to optimize its substrate range for an ODE function. We note that the kinetics of DmJHE for some such substrates (estimated K_M values in the range 10 - 600 µM) are not qualitatively different from the range of values found for some known ODEs from other insects (K_M values in the range 1 µM – 10 mM), suggesting a duplicate expressed in the antennae could then have had immediate utility (Durand et al., 2011; N. Durand et al., 2010; He et al., 2014; Ishida and Leal, 2008, 2005).

An alternative model for the crucial early step in the neofunctionalization of DmJHEdup might propose a structural mutation which obliterated activity against JHIII but retained potentially useful promiscuous activities as a template for further evolution. We cannot discount this possibility and indeed our data suggest changes such as L299/T309 could affect such a transition. A possible problem with this scenario however may be that without a concomitant change in expression profile to produce the enzyme in antennae, it is not clear what immediate utility the new mutation would confer that would prevent the gene from decaying to pseudogene status and retain it as a template for further evolution.

Two broad functional classes of ODEs, specific and general ODEs, have been recognized in insect antennae and other sensory organs (Leal, 2013). Specific ODEs only act on specific substrates and play a dynamic role in refreshing the sensory system to continually respond to new incoming signals (Leal, 2013). The best characterized examples act on sex pheromones, where males in flight must be able to react to changes in the concentration of a female's pheromone plume on a millisecond scale (N. Durand et al., 2010; Nicolas Durand et al., 2010; Vogt et al., 1985). The ApolPDE enzyme of *Antheraea polyphemus*, for example, has a specificity constant for its substrate, E6Z11-16:acetate, of 1.00 x 10^8 M^{-1} s^{-1}, with a K_M of 1.27 µM and k_{cat} of 127 s^{-1} (Ishida and Leal, 2005). On the other hand, general ODEs act on a broader range of substrates and do not have kinetics optimized for a specific substrate. As an example, DmEST6 is proposed to be a general ODE that enables the fly to locate foods emitting various volatile short-mid chain esters and it has specificity constants for these esters estimated to be in the range of 10^5 – 10^6 M^{-1} s^{-1}, with K_M values in the range of 121 – 880 µM (Younus et al., 2017). While our data show

DmJHEdup prefers mid-long chain esters, we can now also see it has evolved kinetics for those esters that are in a very comparable range. In combination with the electrophysiological and behavioral data of Steiner et al (2017), our results show DmJHEdup has effectively neofunctionalized to an ODE role with a complementary set of food esters (Steiner et al., 2017).

Funding:

This work was supported by the Australian Research Council (Future Fellowship to C.J.J.; FT140101059), Australian Science and Industry Endowment Fund (C.J.J.; PF14-099), and by an Australian Government Research Training Program (RTP) Scholarship (D.H.H.).

Acknowledgements:

We thank Mr. Greg Dojchinov and Dr. Shoko Okada for their help and advice with the GC-MS equipment and Dr. Carol J. Hartley for her help with the LC/MS equipment.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Bai, H., Ramaseshadri, P., Palli, S.R., 2007. Identification and characterization of juvenile hormone esterase gene from the yellow fever mosquito, Aedes aegypti. Insect Biochem. Mol. Biol. 37, 829–837.

Barry, J., Wang, S., Wilson, T.G., 2008. Overexpression of Methoprene-tolerant, a Drosophila melanogaster gene that is critical for juvenile hormone action and insecticide resistance. Insect Biochem. Mol. Biol. 38, 346–353.

Benkert, P., Tosatto, S.C.E., Schomburg, D., 2008. QMEAN: A comprehensive scoring function for model quality assessment. Proteins Struct. Funct. Bioinforma. 71, 261–277.

Bianco, G., Forli, S., Goodsell, D.S., Olson, A.J., 2016. Covalent docking using autodock: Two-point attractor and flexible side chain methods. Protein Sci. 25, 295–301.

Campbell, P.M., Harcourt, R.L., Crone, E.J., Claudianos, C., Hammock, B.D., Russell, R.J., Oakeshott, J.G., 2001. Identification of a juvenile hormone esterase gene by matching its peptide mass fingerprint with a sequence from the Drosophila genome project. Insect Biochem. Mol. Biol. 31, 513–520.

Campbell, P.M., Oakeshott, J.G., Healy, M.J., 1998. Purification and kinetic characterisation of juvenile hormone esterase from Drosophila melanogaster. Insect Biochem. Mol. Biol. 28, 501–515.

Capella-Gutierrez, S., Silla-Martinez, J.M., Gabaldon, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973.

Carroll, S.B., 2005. Evolution at two levels: on genes and form. PLoS Biol. 3, e245.

Chertemps, T., François, A., Durand, N., Rosell, G., Dekker, T., Lucas, P., Maïbèche-Coisne, M., 2012. A carboxylesterase, Esterase-6, modulates sensory physiological and behavioral response dynamics to pheromone in Drosophila. BMC Biol. 10, 56.

Chung, H., Bogwitz, M.R., McCart, C., Andrianopoulos, A., Ffrench-Constant, R.H., Batterham, P., Daborn, P.J., 2007. Cis-regulatory elements in the Accord retrotransposon result in tissue-specific expression of the Drosophila melanogaster insecticide resistance gene Cyp6g1. Genetics 175, 1071–1077.

Copley, S.D., 2010. Evolution and the enzyme, in: Barton, D., Meth-Cohn, D. (Eds.), Comprehensive Natural Products Chemistry II. Elsevier, pp. 9–46.

Cornette, R., Gotoh, H., Koshikawa, S., Miura, T., 2008. Juvenile hormone titers and caste differentiation in the damp-wood termite Hodotermopsis sjostedti (Isoptera, Termopsidae). J. Insect Physiol. 54, 922–930.

Cornish-Bowden, A., 1995. Fundamentals of enzyme kinetics. Portland Press, London.

Correy, G.J., Carr, P.D., Meirelles, T., Mabbitt, P.D., Fraser, N.J., Weik, M., Jackson, C.J., 2016. Mapping the accessible conformational landscape of an insect carboxylesterase using conformational ensemble analysis and kinetic crystallography. Structure 24, 977–987.

Crone, E.J., Sutherland, T.D., Campbell, P.M., Coppin, C.W., Russell, R.J., Oakeshott, J.G., 2007. Only one esterase of Drosophila melanogaster is likely to degrade juvenile hormone in vivo. Insect Biochem. Mol. Biol. 37, 540–549.

Cusson, M., Yagi, K.J., Ding, Q., Duve, H., Thorpe, A., McNeil, J.N., Tobe, S.S., 1991. Biosynthesis and release of juvenile hormone and its precursors in insects and crustaceans: The search for a unifying arthropod endocrinology. Insect Biochem. 21, 1–6.

DePristo, M.A., Weinreich, D.M., Hartl, D.L., 2005. Missense meanderings in sequence space: a biophysical view of protein evolution. Nat. Rev. Genet. 6, 678–687.

Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y., Liang, J., 2006. CASTp: Computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. Nucleic Acids Res. 34, 116–118.

Durand, N., Carot-Sans, G., Bozzolan, F., Rosell, G., Siaussat, D., Debernard, S., Chertemps, T., Maïbèche-Coisne, M., 2011. Degradation of pheromone and plant volatile components by a same odorant-degrading enzyme in the cotton leafworm, Spodoptera littoralis. PLoS One 6, e29147.

Durand, N., Carot-Sans, G., Chertemps, T., Bozzolan, F., Party, V., Renou, M., Debernard, S., Rosell, G., Maïbèche-Coisne, M., 2010. Characterization of an antennal carboxylesterase from the pest moth Spodoptera littoralis degrading a host plant odorant. PLoS One 5, e15026.

Durand, N., Carot-Sans, G., Chertemps, T., Montagné, N., Jacquin-Joly, E., Debernard, S., Maïbèche-Coisne, M., 2010. A diversity of putative carboxylesterases are expressed in the antennae of the noctuid moth Spodoptera littoralis. Insect Mol. Biol. 19, 87–97.

El-Sheikh, E.A., 2015. Characterization and kinetics of juvenile hormone esterase from Spodoptera littoralis (Boisd.) and S. frugiperda (J. E. Smith) (Lepidoptera: Noctuidae). Egypt. J. Biol. Pest Control 25, 625–630.

Elayidam, U.G., Muraleedharen, D., 2008. Identification and partial characterization of juvenile hormone esterase from cotton pest Dysdercus cingulatus. Indian J. Biochem. Biophys. 45, 121–125.

Gilbert, L.I., Granger, N.A., Roe, R.M., n.d. The juvenile hormones: historical facts and speculations on future research directions. Insect Biochem. Mol. Biol. 30, 617–644.

Hahn, M.W., Han, M. V., Han, S.-G., 2007. Gene family evolution across 12 Drosophila genomes. PLoS Genet. 3, e197.

Hajós, J.P., Vermunt, A.M., Zuidema, D., Kulcsár, P., Varjas, L., de Kort, C.A., Závodszky, P., Vlak, J.M., 1999. Dissecting insect development: baculovirus-mediated gene silencing in insects. Insect Mol. Biol. 8, 539–544.

Hammock, B.D., 1985. Regulation of juvenile hormone titer: Degradation, in: Kerkut, G.A., Gilbert, L.I. (Eds.), Comprehensive Insect Physiology, Biochemistry, and Pharmacology. Pergamon Press, New York, NY, USA, pp. 431–472.

Hammock, B.D., Bonning, B.C., Possee, R.D., Hanzlik, T.N., Maeda, S., 1990. Expression and effects of the juvenile hormone esterase in a baculovirus vector. Nature 344, 458–461.

He, P., Zhang, J., Li, Z.-Q., Zhang, Y.-N., Yang, K., Dong, S.-L., 2014. Functional characterization of an antennal esterase from the noctuid moth, Spodoptera exigua. Arch. Insect Biochem. Physiol. 86, 85–99.

Hinton, A.C., Hammock, B.D., 2003. Juvenile hormone esterase (JHE) from Tenebrio molitor: full-length cDNA sequence, in vitro expression, and characterization of the recombinant protein. Insect Biochem. Mol. Biol. 33, 477–487.

Hoekstra, H.E., Coyne, J.A., 2007. The locus of evolution: evo devo and the genetics of adaptation. Evolution (N. Y). 61, 995–1016.

Ichikawa, A., Ono, H., Furuta, K., Shiotsuki, T., Shinoda, T., 2007. Enantioselective separation of racemic juvenile hormone III by normal-phase high-performance liquid chromatography and preparation of [2H3]juvenile hormone III as an internal standard for liquid chromatography–mass spectrometry quantification. J. Chromatogr. A 1161, 252–260.

Ishida, Y., Leal, W.S., 2008. Chiral discrimination of the Japanese beetle sex pheromone and a behavioral antagonist by a pheromone-degrading enzyme. Proc. Natl. Acad. Sci. U. S. A. 105, 9076–9080.

Ishida, Y., Leal, W.S., 2005. Rapid inactivation of a moth pheromone. Proc. Natl. Acad. Sci. U. S. A. 102, 14075–14079.

Jeong, S., Rokas, A., Carroll, S.B., 2006. Regulation of body pigmentation by the Abdominal-B Hox protein and its gain and loss in Drosophila evolution. Cell 125, 1387–1399.

Jindra, M., Palli, S.R., Riddiford, L.M., 2013. The juvenile hormone signaling pathway in insect development. Annu. Rev. Entomol. 58, 181–204.

Juneja, P., Quinn, A., Jiggins, F.M., 2016. Latitudinal clines in gene expression and cis-regulatory element variation in Drosophila melanogaster. BMC Genomics 17, 981.

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589.

Kamita, S.G., Hammock, B.D., 2010. Juvenile hormone esterase: biochemistry and structure. J. Pestic. Sci. 35, 265–274.

Kamita, S.G., Hinton, A.C., Wheelock, C.E., Wogulis, M.D., Wilson, D.K., Wolf, N.M., Stok, J.E., Hock, B., Hammock, B.D., 2003. Juvenile hormone (JH) esterase: why are you so JH specific? Insect Biochem. Mol. Biol. 33, 1261–1273.

Kamita, S.G., Samra, A.I., Liu, J.-Y., Cornel, A.J., Hammock, B.D., 2011. Juvenile hormone (JH) esterase of the mosquito Culex quinquefasciatus is not a target of the JH analog insecticide methoprene. PLoS One 6, e28392.

Kamita, S.G., Wogulis, M.D., Law, C.S., Morisseau, C., Tanaka, H., Huang, H., Wilson, D.K., Hammock, B.D., 2010. Function of phenylalanine 259 and threonine 314 within the substrate binding pocket of the juvenile hormone esterase of Manduca sexta. Biochemistry 49, 3733–3742.

Katoh, K., Kuma, K., Toh, H., Miyata, T., 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res. 33, 511–518.

Kidwell, M.G., Lisch, D., 1997. Transposable elements as sources of variation in animals and plants. Proc. Natl. Acad. Sci. U. S. A. 94, 7704–7711.

Kim, D.E., Chivian, D., Baker, D., 2004. Protein structure prediction and analysis using the Robetta server. Nucleic Acids Res. 32, W526–W531.

Kondrashov, F.A., 2012. Gene duplication as a mechanism of genomic adaptation to a changing

environment. Proc. R. Soc. B 279, 5048–5057.

Kontogiannatos, D., Michail, X., Kourti, A., 2011. Molecular characterization of an ecdysteroid inducible carboxylesterase with GQSCG motif in the corn borer, Sesamia nonagrioides. J. Insect Physiol. 57, 1000–1009.

Kontogiannatos, D., Swevers, L., Maenaka, K., Park, E.Y., Iatrou, K., Kourti, A., 2013. Functional characterization of a juvenile hormone esterase related gene in the moth Sesamia nonagrioides through RNA interference. PLoS One 8, e73834.

Leal, W.S., 2013. Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. Annu. Rev. Entomol. 58, 373–391.

Letunic, I., Bork, P., 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. 44, W242–W245.

Levin, H.L., Moran, J. V., 2011. Dynamic interactions between transposable elements and their hosts. Nat. Rev. Genet. 12, 615–627.

Li, W.-H., Gojobori, T., Nei, M., 1981. Pseudogenes as a paradigm of neutral evolution. Nature 292, 237–239.

Li, X., 2007. Juvenile hormone and methyl farnesoate production in cockroach embryos in relation to dorsal closure and the reproductive modes of different species of cockroaches. Arch. Insect Biochem. Physiol. 66, 159–168.

Li, X., Schuler, M. a, Berenbaum, M.R., 2007. Molecular mechanisms of metabolic resistance to synthetic and natural xenobiotics. Annu. Rev. Entomol. 52, 231–253.

Misof, B., Liu, S., Meusemann, K., Peters, R.S., Donath, A., Mayer, C., Frandsen, P.B., Ware, J., Flouri, T., Beutel, R.G., Niehuis, O., Petersen, M., Izquierdo-Carrasco, F., Wappler, T., Rust, J., Aberer, A.J., Aspöck, U., Aspöck, H., Bartel, D., Blanke, A., Berger, S., Böhm, A., Buckley, T.R., Calcott, B., Chen, J., Friedrich, F., Fukui, M., Fujita, M., Greve, C., Grobe, P., Gu, S., Huang, Y., Jermiin, L.S., Kawahara, A.Y., Krogmann, L., Kubiak, M., Lanfear, R., Letsch, H., Li, Y., Li, Z., Li, J., Lu, H., Machida, R., Mashimo, Y., Kapli, P., McKenna, D.D., Meng, G., Nakagaki, Y., Navarrete-Heredia, J.L., Ott, M., Ou, Y., Pass, G., Podsiadlowski, L., Pohl, H., von Reumont, B.M., Schütte, K., Sekiya, K., Shimizu, S., Slipinski, A., Stamatakis, A., Song, W., Su, X., Szucsich, N.U., Tan, M., Tan, X., Tang, M., Tang, J., Timelthaler, G., Tomizuka, S., Trautwein, M., Tong, X., Uchifune, T., Walzl, M.G., Wiegmann, B.M., Wilbrandt, J., Wipfler, B., Wong, T.K.F., Wu, Q., Wu, G., Xie, Y., Yang, S., Yang, Q., Yeates, D.K., Yoshizawa, K., Zhang, Q., Zhang, R., Zhang, W., Zhang, Y., Zhao, J., Zhou, C., Zhou, L., Ziesmann, T., Zou, S., Li, Y., Xu, X., Zhang, Y., Yang, H., Wang, J., Wang, J., Kjer, K.M., Zhou, X., 2014. Phylogenomics resolves the timing and pattern of insect evolution. Science 346, 763–767.

Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S., Olson, A.J., 2009. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J. Comput. Chem. 30, 2785–2791.

Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274.

Noriega, F.G., 2014. Juvenile hormone biosynthesis in insects: What is new, what do we know, and

what questions remain? Int. Sch. Res. Not. 2014, 1–16.

Oakeshott, J.G., Claudianos, C., Campbell, P.M., Newcomb, R.D., Russell, R.J., 2005. Biochemical genetics and genomics of insect esterases, in: Comprehensive Molecular Insect Science. Elsevier Ltd, pp. 309–381.

Oakeshott, J.G., Johnson, R.M., Berenbaum, M.R., Ranson, H., Cristino, A.S., Claudianos, C., 2010. Metabolic enzymes associated with xenobiotic and chemosensory responses in Nasonia vitripennis. Insect Mol. Biol. 19, 147–163.

Oakeshott, J.G., van Papenrecht, E.A., Boyce, T.M., Healy, M.J., Russell, R.J., 1993. Evolutionary genetics of Drosophila esterases. Genetica 90, 239–268.

Podlaha, O., Zhang, J., 2010. Pseudogenes and their evolution, in: Encyclopedia of Life Sciences. John Wiley & Sons, Ltd, Chichester, UK, pp. 1–8.

Robinson, G.E., Strambi, C., Strambi, A., Feldlaufer, M.F., 1991. Comparison of juvenile hormone and ecdysteroid haemolymph titres in adult worker and queen honey bees (Apis mellifera). J. Insect Physiol. 37, 929–935.

Sadd, B.M., Barribeau, S.M., Bloch, G., de Graaf, D.C., Dearden, P., Elsik, C.G., Gadau, J., Grimmelikhuijzen, C.J.P., Hasselmann, M., Lozier, J.D., Robertson, H.M., Smagghe, G., Stolle, E., Van Vaerenbergh, M., Waterhouse, R.M., Bornberg-Bauer, E., Klasberg, S., Bennett, A.K., Câmara, F., Guigó, R., Hoff, K., Mariotti, M., Munoz-Torres, M., Murphy, T., Santesmasses, D., Amdam, G. V, Beckers, M., Beye, M., Biewer, M., Bitondi, M.M.G., Blaxter, M.L., Bourke, A.F.G., Brown, M.J.F., Buechel, S.D., Cameron, R., Cappelle, K., Carolan, J.C., Christiaens, O., Ciborowski, K.L., Clarke, D.F., Colgan, T.J., Collins, D.H., Cridge, A.G., Dalmay, T., Dreier, S., du Plessis, L., Duncan, E., Erler, S., Evans, J., Falcon, T., Flores, K., Freitas, F.C.P., Fuchikawa, T., Gempe, T., Hartfelder, K., Hauser, F., Helbing, S., Humann, F.C., Irvine, F., Jermiin, L.S., Johnson, C.E., Johnson, R.M., Jones, A.K., Kadowaki, T., Kidner, J.H., Koch, V., Köhler, A., Kraus, F.B., Lattorff, H.M.G., Leask, M., Lockett, G.A., Mallon, E.B., Antonio, D.S.M., Marxer, M., Meeus, I., Moritz, R.F.A., Nair, A., Näpflin, K., Nissen, I., Niu, J., Nunes, F.M.F., Oakeshott, J.G., Osborne, A., Otte, M., Pinheiro, D.G., Rossié, N., Rueppell, O., Santos, C.G., Schmid-Hempel, R., Schmitt, B.D., Schulte, C., Simões, Z.L.P., Soares, M.P.M., Swevers, L., Winnebeck, E.C., Wolschin, F., Yu, N., Zdobnov, E.M., Aqrawi, P.K., Blankenburg, K.P., Coyle, M., Francisco, L., Hernandez, A.G., Holder, M., Hudson, M.E., Jackson, L., Jayaseelan, J., Joshi, V., Kovar, C., Lee, S.L., Mata, R., Mathew, T., Newsham, I.F., Ngo, R., Okwuonu, G., Pham, C., Pu, L.-L., Saada, N., Santibanez, J., Simmons, D., Thornton, R., Venkat, A., Walden, K.K.O., Wu, Y.-Q., Debyser, G., Devreese, B., Asher, C., Blommaert, J., Chipman, A.D., Chittka, L., Fouks, B., Liu, J., O'Neill, M.P., Sumner, S., Puiu, D., Qu, J., Salzberg, S.L., Scherer, S.E., Muzny, D.M., Richards, S., Robinson, G.E., Gibbs, R.A., Schmid-Hempel, P., Worley, K.C., 2015. The genomes of two key bumblebee species with primitive eusocial organization. Genome Biol. 16, 76.

Schlenke, T.A., Begun, D.J., 2004. Strong selective sweep associated with a transposon insertion in Drosophila simulans. Proc. Natl. Acad. Sci. U. S. A. 101, 1626–1631.

Schrodinger LLC, 2010. The PyMOL Molecular Graphics System, Version 1.3r1.

Segel, I.H., 1993. Enzyme kinetics: behavior and analysis of rapid equilibrium and steady state enzyme

systems. John Wiley & Sons.

Smykal, V., Bajgar, A., Provaznik, J., Fexova, S., Buricova, M., Takaki, K., Hodkova, M., Jindra, M., Dolezel, D., 2014. Juvenile hormone signaling during reproduction and development of the linden bug, Pyrrhocoris apterus. Insect Biochem. Mol. Biol. 45, 69–76.

Steiner, C., Bozzolan, F., Montagné, N., Maïbèche, M., Chertemps, T., 2017. Neofunctionalization of "juvenile hormone esterase duplication" in Drosophila as an odorant-degrading enzyme towards food odorants. Sci. Rep. 7, 12629.

Tangwancharoen, S., Moy, G.W., Burton, R.S., 2018. Multiple modes of adaptation: regulatory and structural evolution in a small heat shock protein gene. Mol. Biol. Evol. 35, 2110–2119.

Teal, P.E.A., Jones, D., Jones, G., Torto, B., Nyasembe, V., Borgemeister, C., Alborn, H.T., Kaplan, F., Boucias, D., Lietze, V.U., 2014. Identification of methyl farnesoate from the hemolymph of insects. J. Nat. Prod. 77, 402–405.

Tsubota, T., Minakuchi, C., Nakakura, T., Shinoda, T., Shiotsuki, T., 2010. Molecular characterization of a gene encoding juvenile hormone esterase in the red flour beetle, Tribolium castaneum. Insect Mol. Biol. 19, 527–535.

Tsubota, T., Shimomura, M., Ogura, T., Seino, A., Nakakura, T., Mita, K., Shinoda, T., Shiotsuki, T., 2010. Molecular characterization and functional analysis of novel carboxyl/cholinesterases with GQSAG motif in the silkworm Bombyx mori. Insect Biochem. Mol. Biol. 40, 100–112.

Tsubota, T., Shiotsuki, T., 2010. Genomic analysis of carboxyl/cholinesterase genes in the silkworm Bombyx mori. BMC Genomics 11, 377.

Valaitis, A.P., 1991. Characterization of hemolymph juvenile hormone esterase from Lymantria dispar. Insect Biochem. 21, 583–595.

Vogt, R.G., Riddiford, L.M., Prestwich, G.D., 1985. Kinetic properties of a sex pheromone-degrading enzyme: the sensillar esterase of Antheraea polyphemus. Proc. Natl. Acad. Sci. U. S. A. 82, 8827–8831.

Wilson, T.G., Fabian, J., 1986. A Drosophila melanogaster mutant resistant to a chemical analog of juvenile hormone. Dev. Biol. 118, 190–201.

Wittkopp, P.J., Kalay, G., 2012. Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. Nat. Rev. Genet. 13, 59–69.

Wogulis, M., Wheelock, C.E., Kamita, S.G., Hinton, A.C., Whetstone, P.A., Hammock, B.D., Wilson, D.K., 2006. Structural studies of a potent insect maturation inhibitor bound to the juvenile hormone esterase of Manduca sexta. Biochemistry 45, 4045–4057.

Yeates, D.K., Wiegmann, B.M., 1999. Congruence and controversy: toward a higher-Level phylogeny of Diptera. Annu. Rev. Entomol. 44, 397–428.

Yin, C.M., Takeda, M., Wang, Z.S., 1987. A juvenile hormone analogue, methoprene as a circadian and developmental modulator in Diatraea grandiosella (Pyralidae). J. Insect Physiol. 33, 95–102.

Younus, F., Chertemps, T., Pearce, S.L., Pandey, G., Bozzolan, F., Coppin, C.W., Russell, R.J., Maïbèche-Coisne, M., Oakeshott, J.G., 2014. Identification of candidate odorant degrading gene/enzyme systems in the antennal transcriptome of Drosophila melanogaster. Insect Biochem. Mol. Biol. 53, 30–43.

Younus, F., Fraser, N.J., Coppin, C.W., Liu, J.-W., Correy, G.J., Chertemps, T., Pandey, G., Maïbèche, M., Jackson, C.J., Oakeshott, J.G., 2017. Molecular basis for the behavioral effects of the odorant degrading enzyme Esterase 6 in Drosophila. Sci. Rep. 7, 46188.

Zhu, K., Borrelli, K.W., Greenwood, J.R., Day, T., Abel, R., Farid, R.S., Harder, E., 2014. Docking covalent inhibitors: a parameter free approach to pose prediction and scoring. J. Chem. Inf. Model. 54, 1932–1940.

Zhu, L., Yin, T.-Y., Sun, D., Liu, W., Zhu, F., Lei, C.-L., Wang, X.-P., 2017. Juvenile hormone regulates the differential expression of putative juvenile hormone esterases via methoprene-tolerant in non-diapause-destined and diapause-destined adult female beetle. Gene 627, 373–378.

Supplementary Data

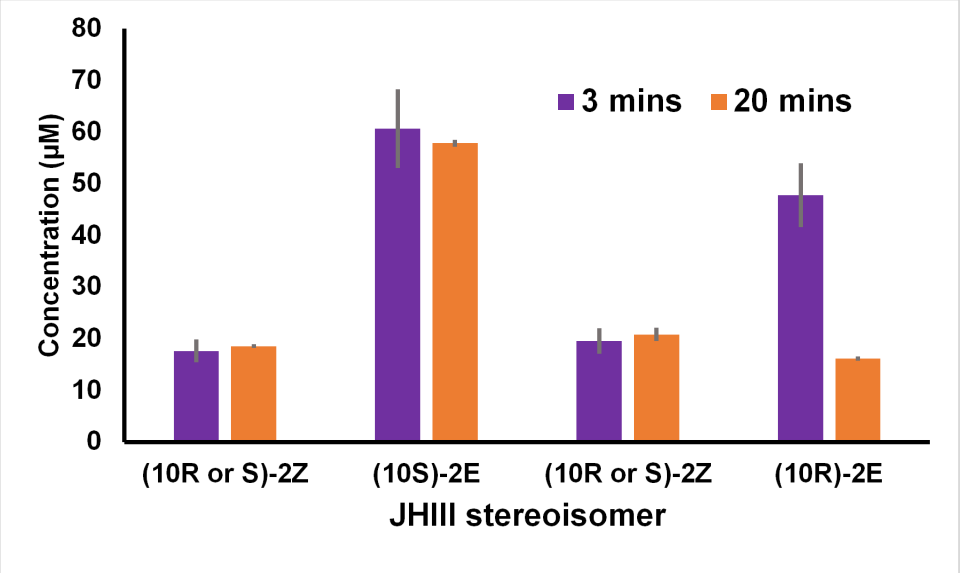


Figure S1. Chiral separation of DmJHE assay with racemic JHIII displaying change over 20 minutes. Stereoisomers are labelled by their differing stereometric configurations. JHIII esters 10R-2Z and 10S-2Z are not distinguished.

Chapter 5. General Discussion

5.1. Insights into the structure, function and evolution of insect CBEs

The work described in this thesis aimed to enhance our understanding of the structure, function and evolution of insect CBEs and to broadly apply this knowledge to the process of enzyme evolution. To do this, I explored two distinct examples of enzyme evolution that have occurred over vastly different periods of time. In Chapters 2 and 3, I focused on the structure, function and evolution of insecticide resistance CBEs, while in Chapter 4, I studied the neofunctionalization of an ODE from a JHE duplicate in *D. melanogaster*.

5.1.1 The structure and function of insecticide resistance CBEs

In Chapter 2, I detail a wide variety of expression trials that resulted in the identification of a number of insecticide resistance CBE candidates for future crystallization. This is a vital first step in enhancing our molecular knowledge of this important family of enzymes. In Chapter 3, I describe the crystallization and characterization of one of these enzymes, Cqestβ2[1], from *C. quinquefasciatus*. This is the first crystal structure of a CBE that acts through insecticide sequestration, the most common form of CBE-mediated insecticide resistance (10, 127, 151, 152). It also adds to our limited knowledge of the structural diversity in the insect CBE family. A comparison between these structures shows that insecticide resistance CBEs have much larger and less specialized substrate binding pockets than CBEs associated with other functions. SSNs reveal that the majority of insecticide resistance CBEs come from a small group of subfamilies related to metabolism. This similarity in sequence may be related to a similarity in both structure and native function between all insecticide resistance CBEs as evidenced by Cqestβ2[1] and LcαE7 (5, 33). I confirmed that Cqestβ2[1] functions through insecticide sequestration and is likely to have a major role in OP resistance and a minor role in carbamate resistance. I explored the sequence diversity of Cqestβ2[1] isoforms through a direct comparison with Cqestβ1, the most divergent isoform. This showed that the 16 amino acid differences between them had little effect on their interactions with insecticides. It also suggests the broad applicability of

insecticides made through target-based inhibitor design with Cqestβ2[1] against its isoforms.

5.1.2. The evolution of insecticide resistance CBEs

Insecticides have been used for a relatively short period of time in terms of eukaryotic evolution (30). Thus, the majority of the mechanisms that have developed provide an example of the early stages of evolution. In the case of Cqestβ2[1], I confirmed that resistance has not evolved through hydrolysis but rather sequestration. In this case, the evolution towards gene amplification, and upregulation of Cqestβ2[1] was selected over structural changes to enhance hydrolysis as found in LcαE7 from *L. cuprina* (5, 163, 164). The equivalent mutation to G137D in LcαE7, which enhances insecticide hydrolysis, was previously shown to have no effect in Cqestβ2[1] (208). This mutation was also shown computationally to rely on one residue, F309, to adopt catalytically productive conformations in LcαE7 (63). I suggest that both the expanded pocket and lack of an equivalent residue in Cqestβ2[1] reduce the viability of the G137D mutation. This exemplifies structural limitations that may prevent the evolution of the OP hydrolysis mechanism in some insecticide resistance CBEs. The apparent evolution of regulatory over structural changes in Cqestβ2[1] may also be related to the fitness costs associated with the loss of Cqestβ2[1] native function by OP hydrolysis enhancing mutations (96, 168, 169). Both Cqestβ2[1] and LcαE7 have a broad specificity, which may be an inherent feature of insecticide resistance CBEs (5, 33, 257). It is likely that the inherent promiscuity of these enzymes and their fortuitous expression profiles have enabled their role as insecticide resistance CBEs.

5.1.3. The structure and function of insect ODEs and JHEs

In Chapter 4, I describe a biochemical and structural comparison between *D. melanogaster* DmJHE and its duplicate, DmJHEdup. As the structure of neither enzyme was known, I utilized homology modelling to determine their likely structures. DmJHE is confirmed to have a strong interaction with JHIII but no interaction with its precursor methyl farnesoate, which also acts as a hormone (258, 259). DmJHE also shows a far broader substrate range than previously thought, which may suggest a similar trait in other insect JHEs. A comparison of the DmJHE model with the structure

of MsJHE reveals key conserved residues that interact with JHIII. These are known to be shared amongst insect JHEs and suggest a level of structural conservation between them (232, 260). I showed that DmJHEdup functions as an ODE with a broad substrate range with a preference for mid-long chain food odorant esters. This substrate range complements DmEST6, the other *D. melanogaster* ODE to be identified. It also indicates that DmJHEdup functions as a general ODE. While the DmJHEdup homology model and DmEST6 structure shared a similar active site entrance, the DmJHEdup binding pocket was much larger consistent with its substrate range. These results enhance our knowledge of both the structure and function of insect ODEs and JHEs.

5.1.4. The neofunctionalization of enzyme duplicates

One of the early principles of enzyme evolution was that all specialized enzymes evolved from ancestral forms with broad specificities (55). Previous phylogenies of insect ODEs and JHEs had each group in distinct monophyletic clades suggesting both emerged from a common ancestral protein, consistent with the above principle. Due to an increase in the number of characterized insect JHEs and ODEs, I was able to generate a phylogeny that reveals more complexity. It instead suggests that the JHE function predates the hemi/holometabolon split in insects and that ODEs with both broad and specific functions have evolved from JHEs through several instances of duplication and diversification. The analysis of DmJHEdup and DmJHE thus exemplifies two key principles in enzyme evolution: the inherent promiscuity of enzymes, even specialist enzymes; and their ability to evolve new functions in the right conditions (54, 252, 261). As JHE is vital to insect development, evolution required a duplication, exemplifying strong negative trade-offs in neofunctionalization (54, 262). Regulatory changes, resulting in the unique expression profile of DmJHEdup, were likely the first step in the evolution of this duplicate towards an ODE function with broad specificity. This is supported by the inherent promiscuity I detected in DmJHE, with which an ancestral form would have immediate utility as an ODE. I suggest a number of structural changes such as T309/L299 that would subsequently lead to the evolution of DmJHEdup's unique activity and loss of JHIII activity. This provides an in-depth analysis of the process of enzyme evolution in *D. melanogaster* and a unique example of the evolution of a generalist from a specialist enzyme.

5.2. Future directions

5.2.1. Insecticide resistance CBEs

I identified a number of other insecticide resistance CBEs with expression sufficient for further characterization and potential crystallization. This would be a useful step in further enhancing our understanding of insecticide resistance CBEs. The crystal structure of Cqestβ2[1] could be used in target-based inhibitor design to create synergistic inhibitors to enhance insecticide effectiveness and help combat resistance (263). The SSN I developed could also be used as a guide to select sequences and generate a better phylogeny of the insect CBE family, which would allow easier classification and functional annotation of insect CBEs.

5.2.2. DmJHE and DmJHEdup

I present the first evidence of inherent promiscuity in insect JHEs. To confirm the broad applicability of this a wider variety of insect JHEs from other insect orders should be tested. It would also be beneficial to determine the structures of both DmJHE and DmJHEdup to confirm that the homology models were good representations of each enzyme and validate the structural suggestions that I made. Similarly, determining the structures of other insect JHEs from distinct orders would better elucidate the structural similarities between them, which could inform the design of better JH analogues for use as insecticides. To further analyze the process of DmJHEdup evolution, ancestral reconstruction and directed evolution could be used. This may provide further insights into the evolution of generalist enzymes from specialists.

References

1. Horgan, D. J., Stoops, J. K., Webb, E. C., and Zerner, B. (1969) Carboxylesterases (EC 3.1.1.). A large-scale purification of pig liver carboxylesterase. *Biochemistry*. **8**, 2000–2006
2. Nussbaumer, C., Hinton, A. C., Schopf, A., Stradner, A., and Hammock, B. D. (2000) Isolation and characterization of juvenile hormone esterase from hemolymph of *Lymantria dispar* by affinity- and by anion-exchange chromatography. *Insect Biochem. Mol. Biol.* **30**, 307–314
3. Pralavorio, M., and Fournier, D. (1992) Drosophila acetylcholinesterase: characterization of different mutants resistant to insecticides. *Biochem. Genet.* **30**, 77–83
4. Ketterman, A. J., Jayawardena, K. G., and Hemingway, J. (1992) Purification and characterization of a carboxylesterase involved in insecticide resistance from the mosquito *Culex quinquefasciatus*. *Biochem. J.* **287**, 355–360
5. Jackson, C. J., Liu, J. W., Carr, P. D., Younus, F., Coppin, C., Meirelles, T., Lethier, M., Pandey, G., Ollis, D. L., Russell, R. J., et al. (2013) Structure and function of an insect α-carboxylesterase (αEsterase7) associated with insecticide resistance. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 10177–10182
6. Casey Laizure, S., Herring, V., Hu, Z., Witbrodt, K., and Parker, R. B. (2013) The role of human carboxylesterases in drug metabolism: have we overlooked their importance? *Pharmacother. J. Hum. Pharmacol. Drug Ther.* **33**, 210–222
7. Xu, G., Zhang, W., Ma, M. K., and McLeod, H. L. (2002) Human carboxylesterase 2 is commonly expressed in tumor tissue and is correlated with activation of irinotecan. *Clin. Cancer Res.* **8**, 2605–2611
8. Sugrue, E., Fraser, N. J., Hopkins, D. H., Carr, P. D., Khurana, J. L., Oakeshott, J. G., Scott, C., and Jackson, C. J. (2015) Evolutionary expansion of the amidohydrolase superfamily in bacteria in response to the synthetic compounds molinate and diuron. *Appl. Environ. Microbiol.* **81**, 2612–2624
9. Duarte, M., Ferreira-da-Silva, F., Lünsdorf, H., Junca, H., Gales, L., Pieper, D. H., and Nunes, O. C. (2011) *Gulosibacter molinativorax* ON4T molinate hydrolase, a novel cobalt-dependent amidohydrolase. *J. Bacteriol.* **193**, 5810–5816
10. Wheelock, C. E., Shan, G., and Ottea, J. (2005) Overview of carboxylesterases

and their role in the metabolism of insecticides. *J. Pestic. Sci.* **30**, 75–83

11. Russell, R. J., Scott, C., Jackson, C. J., Pandey, R., Pandey, G., Taylor, M. C., Coppin, C. W., Liu, J. W., and Oakeshott, J. G. (2011) The evolution of new enzyme function: lessons from xenobiotic metabolizing bacteria versus insecticide-resistant insects. *Evol. Appl.* **4**, 225–248
12. Brown, S. D., and Babbitt, P. C. (2014) New insights about enzyme evolution from large scale studies of sequence and structure relationships. *J. Biol. Chem.* **289**, 30221–30228
13. Liszka, M. J., Clark, M. E., Schneider, E., and Clark, D. S. (2012) Nature versus nurture: developing enzymes that function under extreme conditions. *Annu. Rev. Chem. Biomol. Eng.* **3**, 77–102
14. Miton, C. M., and Tokuriki, N. (2016) How mutational epistasis impairs predictability in protein evolution and design. *Protein Sci.* **25**, 1260–1272
15. Montella, I. R., Schama, R., and Valle, D. (2012) The classification of esterases: an important gene family involved in insecticide resistance - A review. *Mem. Inst. Oswaldo Cruz*. **107**, 437–449
16. Aldridge, W. N. (1953) Serum esterases. 1. Two types of esterase (A and B) hydrolysing p-nitrophenyl acetate, propionate and butyrate, and a method for their determination. *Biochem. J.* **53**, 110–117
17. Aldridge, W. N. (1953) Serum esterases. II. An enzyme hydrolysing diethyl p-nitrophenyl phosphate (E600) and its identity with the A-esterase of mammalian sera. *Biochem. J.* **53**, 117–124
18. Walker, C. H., and Mackness, M. I. (1983) Esterases: problems of identification and classification. *Biochem. Pharmacol.* **32**, 3265–3269
19. Holmes, R. S., and Masters, C. J. (1967) The developmental multiplicity and isoenzyme status of cavian esterases. *Biochim. Biophys. Acta - Enzymol.* **132**, 379–399
20. International Union of Biochemistry (1979) *Enzyme nomenclature 1978. Recommendations of the Nomenclature Committee of the International Union of Biochemistry on the nomenclature and classification of enzymes.*, Academic Press
21. Enzyme nomenclature. Report on the recommendations (1964) of the International Union of Biochemistry on Nomenclature and Classification of Enzymes. (1965) *Science*. **150**, 719–721

22. International Union of Biochemistry and Molecular Biology. (1992) *Enzyme nomenclature 1992 : recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the nomenclature and classification of enzymes*
23. Mander, L. N., and Liu, H. (2010) *Comprehensive natural products II : chemistry and biology*, Elsevier
24. Younus, F., Fraser, N. J., Coppin, C. W., Liu, J.-W., Correy, G. J., Chertemps, T., Pandey, G., Maïbèche, M., Jackson, C. J., and Oakeshott, J. G. (2017) Molecular basis for the behavioral effects of the odorant degrading enzyme Esterase 6 in Drosophila. *Sci. Rep.* **7**, 46188
25. Bass, C., and Field, L. M. (2011) Gene amplification and insecticide resistance. *Pest Manag. Sci.* **67**, 886–890
26. Kondrashov, F. A. (2012) Gene duplication as a mechanism of genomic adaptation to a changing environment. *Proc. R. Soc. B*. **279**, 5048–5057
27. Whalon, M. E., Mota-Sanchez, D. (David), and Hollingworth, R. M. (2008) *Global pesticide resistance in arthropods*, CABI
28. Devonshire, A. L., and Moores, G. D. (1982) A carboxylesterase with broad substrate specificity causes organophosphorus, carbamate and pyrethroid resistance in peach-potato aphids (*Myzus persicae*). *Pestic. Biochem. Physiol.* **18**, 235–246
29. Zouros, E., van Delden, W., Odense, R., and van Dijk, H. (1982) An esterase duplication in Drosophila: differences in expression of duplicate loci within and among related species. *Biochem. Genet.* **20**, 929–942
30. Oakeshott, J. G., Claudianos, C., Campbell, P. M., Newcomb, R. D., and Russell, R. J. (2005) Biochemical genetics and genomics of insect esterases. in *Comprehensive molecular insect science* (Gilbert, L., Iatrou, K., and Gill, S. eds), pp. 309–381, Elsevier
31. Vaughan, A., Rodriguez, M., and Hemingway, J. (1995) The independent gene amplification of electrophoretically indistinguishable B esterases from the insecticide-resistant mosquito *Culex quinquefasciatus*. *Biochem. J.* **305**, 651–658
32. Vaughan, a, and Hemingway, J. (1995) Mosquito carboxylesterase Est alpha 2(1) (A2). Cloning and sequence of the full-length cDNA for a major insecticide resistance gene worldwide in the mosquito *Culex quinquefasciatus*. *J. Biol.*

Chem. **270**, 17044–17049

33. Karunaratne, S. H., Jayawardena, K. G., Hemingway, J., and Ketterman, a J. (1993) Characterization of a B-type esterase involved in insecticide resistance from the mosquito Culex quinquefasciatus. *Biochem. J.* **294**, 575–579
34. Satoh, T., and Hosokawa, M. (2006) Structure, function and regulation of carboxylesterases. *Chem. Biol. Interact.* **162**, 195–211
35. Oakeshott, J. G., Johnson, R. M., Berenbaum, M. R., Ranson, H., Cristino, A. S., and Claudianos, C. (2010) Metabolic enzymes associated with xenobiotic and chemosensory responses in *Nasonia vitripennis*. *Insect Mol. Biol.* **19**, 147–163
36. Sadd, B. M., Barribeau, S. M., Bloch, G., de Graaf, D. C., Dearden, P., Elsik, C. G., Gadau, J., Grimmelikhuijzen, C. J. P., Hasselmann, M., Lozier, J. D., et al. (2015) The genomes of two key bumblebee species with primitive eusocial organization. *Genome Biol.* **16**, 76
37. Teese, M. G., Campbell, P. M., Scott, C., Gordon, K. H. J. J., Southon, A., Hovan, D., Robin, C., Russell, R. J., and Oakeshott, J. G. (2010) Gene identification and proteomic analysis of the esterases of the cotton bollworm, *Helicoverpa armigera*. *Insect Biochem. Mol. Biol.* **40**, 1–16
38. Yu, Q. Y., Lu, C., Li, W. Le, Xiang, Z. H., and Zhang, Z. (2009) Annotation and expression of carboxylesterases in the silkworm, Bombyx mori. *BMC Genomics*. **10**, 553
39. Strode, C., Wondji, C. S., David, J. P., Hawkes, N. J., Lumjuan, N., Nelson, D. R., Drane, D. R., Karunaratne, S. H. P. P., Hemingway, J., Black, W. C., et al. (2008) Genomic analysis of detoxification genes in the mosquito *Aedes aegypti*. *Insect Biochem. Mol. Biol.* **38**, 113–123
40. Lü, F. G., Fu, K. Y., Li, Q., Guo, W. C., Ahmat, T., and Li, G. Q. (2015) Identification of carboxylesterase genes and their expression profiles in the Colorado potato beetle *Leptinotarsa decemlineata* treated with fipronil and cyhalothrin. *Pestic. Biochem. Physiol.* **122**, 86–95
41. Ramsey, J. S., Rider, D. S., Walsh, T. K., De Vos, M., Gordon, K. H. J., Ponnala, L., Macmil, S. L., Roe, B. A., and Jander, G. (2010) Comparative analysis of detoxification enzymes in *Acyrthosiphon pisum* and *Myzus persicae*. *Insect Mol. Biol.* **19**, 155–164
42. Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S.

R., Heger, A., Hetherington, K., Holm, L., Mistry, J., et al. (2014) Pfam: the protein families database. *Nucleic Acids Res.* **42**, D222–D230

43. David, L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Sybille, M., Harel, M., James Remington, S., Silman, I., Schrag, J., et al. (1992) The α/β hydrolase fold. *Protein Eng. Des. Sel.* **5**, 197–211
44. Lenfant, N., Hotelier, T., Velluet, E., Bourne, Y., Marchot, P., and Chatonnet, A. (2012) ESTHER, the database of the α/β-hydrolase fold superfamily of proteins: tools to explore diversity of functions. *Nucleic Acids Res.* **41**, D423–D429
45. Nardini, M., and Dijkstra, B. W. (1999) α/β Hydrolase fold enzymes: the family keeps growing. *Curr. Opin. Struct. Biol.* **9**, 732–737
46. Rauwerdink, A., and Kazlauskas, R. J. (2015) How the same core catalytic machinery catalyzes 17 different reactions: the serine-histidine-aspartate catalytic triad of α/β-hydrolase fold enzymes. *ACS Catal.* **5**, 6153–6176
47. Aranda, J., Cerqueira, N. M. F. S. A., Fernandes, P. A., Roca, M., Tuñón, I., and Ramos, M. J. (2014) The catalytic mechanism of carboxylesterases: a computational study. *Biochemistry*. **53**, 5820–5829
48. Redinbo, M. R., Bencharit, S., and Potter, P. M. (2003) Human carboxylesterase 1: from drug metabolism to drug discovery. *Biochem. Soc. Trans.* **31**, 620–624
49. Satoh, T., Taylor, P., Bosron, W, F., Sanghani, S, P., Hosokawa, M., and La Du, B, N. (2002) Current progress on esterases: from molecular structure to function. *Drug Metab. Dispos.* **30**, 488–493
50. Satoh, T., and Hosokawa, M. (1998) The mammalian carboxylesterases: from molecules to functions. *Annu. Rev. Pharmacol. Toxicol.* **38**, 257–288
51. Hosokawa, M., Endo, T., Fujisawa, M., Hara, S., Iwata, N., Sato, Y., and Satoh, T. (1995) Interindividual variation in carboxylesterase levels in human liver microsomes. *Drug Metab. Dispos.* **23**, 1022–1027
52. Sogorb, M. A., and Vilanova, E. (2002) Enzymes involved in the detoxification of organophosphorus, carbamate and pyrethroid insecticides through hydrolysis. *Toxicol. Lett.* **128**, 215–228
53. Dodson, G., and Wlodawer, A. (1998) Catalytic triads and their relatives. *Trends Biochem. Sci.* **23**, 347–352
54. Tawfik, O. K. and D. S. (2010) Enzyme promiscuity: a mechanistic and evolutionary perspective. *Annu. Rev. Biochem.* **79**, 471–505
55. Jensen, R. A. (1976) Enzyme recruitment in evolution of new function. *Annu.*

Rev. Microbiol. **30**, 409–425

56. Aharoni, A., Gaidukov, L., Khersonsky, O., Gould, S. M., Roodveldt, C., and Tawfik, D. S. (2005) The “evolvability” of promiscuous protein functions. *Nat. Genet.* **37**, 73–76
57. O’Brien, P. J., and Herschlag, D. (1999) Catalytic promiscuity and the evolution of new enzymatic activities. *Chem. Biol.* **6**, R91–R105
58. Copley, S. D. (2003) Enzymes with extra talents: moonlighting functions and catalytic promiscuity. *Curr. Opin. Chem. Biol.* **7**, 265–272
59. Patrick, W. M., Quandt, E. M., Swartzlander, D. B., and Matsumura, I. (2007) Multicopy suppression underpins metabolic evolvability. *Mol. Biol. Evol.* **24**, 2716–2722
60. Tokuriki, N., and Tawfik, D. S. (2009) Protein dynamism and evolvability. *Science*. **324**, 203–207
61. Clifton, B. E., and Jackson, C. J. (2016) Ancestral protein reconstruction yields insights into adaptive evolution of binding specificity in solute-binding proteins. *Cell Chem. Biol.* **23**, 236–245
62. Campbell, E., Kaltenbach, M., Correy, G. J., Carr, P. D., Porebski, B. T., Livingstone, E. K., Afriat-Jurnou, L., Buckle, A. M., Weik, M., Hollfelder, F., et al. (2016) The role of protein dynamics in the evolution of new enzyme function. *Nat. Chem. Biol.* **12**, 944–950
63. Mabbitt, P. D., Correy, G. J., Meirelles, T., Fraser, N. J., Coote, M. L., and Jackson, C. J. (2016) Conformational disorganization within the active site of a recently evolved organophosphate hydrolase limits its catalytic efficiency. *Biochemistry*. **55**, 1408–1417
64. Copley, S. D. (2010) Evolution and the enzyme. in *Comprehensive Natural Products Chemistry II* (Barton, D., and Meth-Cohn, D. eds), pp. 9–46, Elsevier
65. O’Loughlin, T. L., Greene, D. N., and Matsumura, I. (2006) Diversification and specialization of HIV protease function during in vitro evolution. *Mol. Biol. Evol.* **23**, 764–772
66. Ran, N., Draths, K. M., and Frost, J. W. (2004) Creation of a shikimate pathway variant. *J. Am. Chem. Soc.* **126**, 6856–6857
67. Varadarajan, N., Gam, J., Olsen, M. J., Georgiou, G., and Iverson, B. L. (2005) Engineering of protease variants exhibiting high catalytic activity and exquisite substrate selectivity. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 6855–6860

68. Khersonsky, O., Roodveldt, C., and Tawfik, D. S. (2006) Enzyme promiscuity: evolutionary and mechanistic aspects. *Curr. Opin. Chem. Biol.* **10**, 498–508
69. Ju, K. S., Parales, J. V., and Parales, R. E. (2009) Reconstructing the evolutionary history of nitrotoluene detection in the transcriptional regulator NtdR. *Mol. Microbiol.* **74**, 826–843
70. Kivisaar, M. (2009) Degradation of nitroaromatic compounds: a model to study evolution of metabolic pathways. *Mol. Microbiol.* **74**, 777–781
71. Levin, K. B., Dym, O., Albeck, S., Magdassi, S., Keeble, A. H., Kleanthous, C., and Tawfik, D. S. (2009) Following evolutionary paths to protein-protein interactions with high affinity and selectivity. *Nat. Struct. Mol. Biol.* **16**, 1049–1055
72. Thornton, J. W., Need, E., and Crews, D. (2003) Resurrecting the ancestral steroid receptor: ancient origin of estrogen signaling. *Science (80-.).* **301**, 1714–1717
73. Watts, K. T., Mijts, B. N., Lee, P. C., Manning, A. J., and Schmidt-Dannert, C. (2006) Discovery of a substrate selectivity switch in tyrosine ammonia-lyase, a member of the aromatic amino acid lyase family. *Chem. Biol.* **13**, 1317–1326
74. Wang, X., Minasov, G., and Shoichet, B. K. (2002) Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs. *J. Mol. Biol.* **320**, 85–95
75. DePristo, M. A., Weinreich, D. M., and Hartl, D. L. (2005) Missense meanderings in sequence space: a biophysical view of protein evolution. *Nat. Rev. Genet.* **6**, 678–687
76. Bloom, J. D., Silberg, J. J., Wilke, C. O., Drummond, D. A., Adami, C., and Arnold, F. H. (2005) Thermodynamic prediction of protein neutrality. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 606–611
77. Camps, M., Herman, A., Loh, E., and Loeb, L. A. (2007) Genetic constraints on protein evolution. *Crit. Rev. Biochem. Mol. Biol.* **42**, 313–326
78. Tokuriki, N., and Tawfik, D. S. (2009) Stability effects of mutations and protein evolvability. *Curr. Opin. Struct. Biol.* **19**, 596–604
79. Poelwijk, F. J., Kiviet, D. J., Weinreich, D. M., and Tans, S. J. (2007) Empirical fitness landscapes reveal accessible evolutionary paths. *Nature*. **445**, 383–386
80. Benning, M. M., Kuo, J. M., Raushel, F. M., and Holden, H. M. (1995) Three-dimensional structure of the binuclear metal center of phosphotriesterase.

Biochemistry. **34**, 7973–7978

81. Cygler, M., Schrag, J. D., and Ergan, F. (1992) Advances in structural understanding of lipases. *Biotechnol. Genet. Eng. Rev.* **10**, 143–184
82. Doolittle, R. F., Feng, D. F., Tsang, S., Cho, G., and Little, E. (1996) Determining divergence times of the major kingdoms of living organisms with a protein clock. *Science (80-.).* **271**, 470–477
83. Hemilä, H., Koivula, T. T., and Palva, I. (1994) Hormone-sensitive lipase is closely related to several bacterial proteins, and distantly related to acetylcholinesterase and lipoprotein lipase: Identification of a superfamily of esterases and lipases. *Biochim. Biophys. Acta - Lipids Lipid Metab.* **1210**, 249–253
84. Rauwerdink, A. M., and Kazlauskas, R. J. (2015) How the same core catalytic machinery catalyzes seventeen different reactions: the Ser-His-Asp catalytic triad of α/β-hydrolase fold enzymes. *ACS Catal.* **5**, 6153–6176
85. Lenfant, N., Hotelier, T., Bourne, Y., Marchot, P., and Chatonnet, A. (2013) Proteins with an alpha/beta hydrolase fold: Relationships between subfamilies in an ever-growing superfamily. *Chem. Biol. Interact.* **203**, 266–268
86. Oakeshott, J. G., Claudianos, C., Russell, R. J., and Robin, G. C. (1999) Carboxyl/cholinesterases: a case study of the evolution of a successful multigene family. *Bioessays*. **21**, 1031–1042
87. Putterill, J. J., Plummer, K. M., Newcomb, R. D., and Marshall, S. D. G. (2003) The Carboxylesterase Gene Family from Arabidopsis thaliana. *J. Mol. Evol.* **57**, 487–500
88. Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J. D., Jacq, C., Johnston, M., et al. (1996) Life with 6000 Genes. *Science (80-.).* **274**, 546–567
89. Lomolino, G., Lante, A., Rizzi, C., Spettoli, P., and Curioni, A. (2005) Comparison of esterase patterns of three yeast strains as obtained with different synthetic substrates. *J. Inst. Brew.* **111**, 234–236
90. Horsted, M. W., Dey, E. S., Holmberg, S., and Kielland-Brandt, M. C. (1998) A novel esterase from *Saccharomyces carlsbergensis*, a possible function for the yeast TIP1 gene. *Yeast*. **14**, 793–803
91. Hahn, M. W., Han, M. V., and Han, S. G. (2007) Gene family evolution across 12 Drosophila genomes. *PLoS Genet.* **3**, e197

92. Robin, C., Russell, R. J., Medveczky, K. M., and Oakeshott, J. G. (1996) Duplication and divergence of the genes of the α-esterase cluster of *Drosophila melanogaster*. *J. Mol. Evol.* **43**, 241–252
93. Powell, J. R., and DeSalle, R. (1995) Drosophila molecular phylogenies and their uses. in *Evolutionary Biology*, pp. 87–138, Springer US, Boston, MA
94. Karunaratne, S. H. P. P., and Hemingway, J. (2001) Malathion resistance and prevalence of the malathion carboxylesterase mechanism in populations of mosquito vectors of disease in Sri Lanka. *Bull. World Health Organ.* **79**, 1060–1064
95. Ranson, H., Claudianos, C., Ortelli, F., Abgrall, C., Hemingway, J., Sharakhova, M. V, Unger, M. F., Collins, F. H., and Feyereisen, R. (2002) Evolution of supergene families associated with insecticide resistance. *Science*. **298**, 179–181
96. Cui, F., Li, M. X., Chang, H. J., Mao, Y., Zhang, H. Y., Lu, L. X., Yan, S. G., Lang, M. L., Liu, L., and Qiao, C. L. (2015) Carboxylesterase-mediated insecticide resistance: quantitative increase induces broader metabolic resistance than qualitative change. *Pestic. Biochem. Physiol.* **121**, 88–96
97. Hemingway, J., and Ranson, H. (2000) Insecticide resistance in insect vectors of human disease. *Annu Rev Entomol*. **45**, 371–391
98. Hardy, M. C. (2014) Resistance is not futile: it shapes insecticide discovery. *Insects*. **5**, 227–242
99. Rivero, A., Vézilier, J., Weill, M., Read, A. F., and Gandon, S. (2010) Insecticide control of vector-borne diseases: when is insecticide resistance a problem? *PLoS Pathog.* **6**, 5–6
100. ffrench-Constant, R. H., and Bass, C. (2017) Does resistance really carry a fitness cost? *Curr. Opin. Insect Sci.* **21**, 39–46
101. Labbe, P., Lenormand, T., and Raymond, M. (2005) On the worldwide spread of an insecticide resistance gene: a role for local selection. *J. Evol. Biol.* **18**, 1471–1484
102. Sparks, T. C., and Nauen, R. (2015) IRAC: Mode of action classification and insecticide resistance management. *Pestic. Biochem. Physiol.* **121**, 122–128
103. Jeschke, P., Nauen, R., Schindler, M., and Elbert, A. (2011) Overview of the status and global strategy for neonicotinoids. *J. Agric. Food Chem.* **59**, 2897–2908

104. Zhu, Y., Loso, M. R., Watson, G. B., Sparks, T. C., Rogers, R. B., Huang, J. X., Gerwick, B. C., Babcock, J. M., Kelley, D., Hegde, V. B., et al. (2011) Discovery and characterization of sulfoxaflor, a novel insecticide targeting sap-feeding pests. *J. Agric. Food Chem.* **59**, 2950–2957
105. Zaim, M., and Guillet, P. (2002) Alternative insecticides: an urgent need. *Trends Parasitol.* **18**, 161–163
106. Sparks, T. C. (2013) Insecticide discovery: an evaluation and analysis. *Pestic. Biochem. Physiol.* **107**, 8–17
107. Alavanja, M. C. R. (2009) Pesticides use and exposure extensive worldwide. *Rev. Environ. Health*. **24**, 303–309
108. Atwood, D., and Paisley-Jones, C. (2017) Pesticides Industry Sales and Usage 2008 - 2012 Market Estimates. *United States Environ. Prot. Agency Washington, DC, USA*
109. Li, X., Schuler, M. A., and Berenbaum, M. R. (2007) Molecular mechanisms of metabolic resistance to synthetic and natural xenobiotics. *Annu. Rev. Entomol.* **52**, 231–253
110. Mileson, B. E., Chambers, J. E., Chen, W. L., Dettbarn, W., Ehrich, M., Eldefrawi, A. T., Gaylor, D. W., Hamernik, K., Hodgson, E., Karczmar, A. G., et al. (1998) Common mechanism of toxicity: a case study of organophosphorus pesticides. *Toxicol. Sci.* **41**, 8–20
111. Kousba, A. A., Sultatos, L. G., Poet, T. S., and Timchalk, C. (2004) Comparison of chlorpyrifos-oxon and paraoxon acetylcholinesterase inhibition dynamics: potential role of a peripheral binding site. *Toxicol. Sci.* **80**, 239–248
112. Forsberg, A., and Puu, G. (1984) Kinetics for the inhibition of acetylcholinesterase from the electric eel by some organophosphates and carbamates. *Eur. J. Biochem.* **140**, 153–156
113. Hart, G. J., and O'Brien, R. D. (1974) Stopped-flow studies of the inhibition of acetylcholinesterase by organophosphates in the presence of substrate. *Pestic. Biochem. Physiol.* **4**, 239–244
114. Colović, M. B., Krstić, D. Z., Lazarević-Pašti, T. D., Bondžić, A. M., and Vasić, V. M. (2013) Acetylcholinesterase inhibitors: pharmacology and toxicology. *Curr. Neuropharmacol.* **11**, 315–335
115. Harel, M., Kryger, G., Rosenberry, T. L., Mallender, W. D., Lewis, T., Fletcher, R. J., Guss, J. M., Silman, I., and Sussman, J. L. (2000) Three-dimensional

structures of *Drosophila melanogaster* acetylcholinesterase and of its complexes with two potent inhibitors. *Protein Sci.* **9**, 1063–1072

116. Zhu, K. Y., and Brindley, W. A. (1992) Enzymological and inhibitory properties of acetylcholinesterase purified from *Lygus hesperus* knight (Hemiptera:Miridae). *Insect Biochem. Mol. Biol.* **22**, 245–251
117. Chen, Z., Newcomb, R., Forbes, E., McKenzie, J., and Batterham, P. (2001) The acetylcholinesterase gene and organophosphorus resistance in the Australian sheep blowfly, *Lucilia cuprina*. *Insect Biochem. Mol. Biol.* **31**, 805–816
118. Sultatos, L. G., and Murphy, S. D. (1983) Kinetic analyses of the microsomal biotransformation of the phosphorothioate insecticides chlorpyrifos and parathion. *Toxicol. Sci.* **3**, 16–21
119. Walz, I., and Schwack, W. (2007) Multienzyme inhibition assay for residue analysis of insecticidal organophosphates and carbamates. *J. Agric. Food Chem.* **55**, 10563–10571
120. Ghazala, Mahboob, S., Ahmad, L., Sultana, S., Alghanim, K., Al-Misned, F., and Ahmad, Z. (2014) Fish cholinesterases as biomarkers of sublethal effects of organophosphorus and carbamates in tissues of *Labeo rohita*. *J. Biochem. Mol. Toxicol.* **28**, 137–142
121. Millard, C. B., Kryger, G., Ordentlich, A., Greenblatt, H. M., Harel, M., Raves, M. L., Segall, Y., Barak, D., Shafferman, A., Silman, I., et al. (1999) Crystal structures of aged phosphonylated acetylcholinesterase: nerve agent reaction products at the atomic Level. *Biochemistry*. **38**, 7032–7039
122. Nachon, F., Asojo, O. A., Borgstahl, G. E. O., Masson, P., and Lockridge, O. (2005) Role of water in aging of human butyrylcholinesterase inhibited by echothiophate: the crystal structure suggests two alternative mechanisms of aging. *Biochemistry*. **44**, 1154–1162
123. Li, H., Schopfer, L. M., Nachon, F., Froment, M. T., Masson, P., and Lockridge, O. (2007) Aging pathways for organophosphate-inhibited human butyrylcholinesterase, including novel pathways for isomalathion, resolved by mass spectrometry. *Toxicol. Sci.* **100**, 136–145
124. Shafferman, A., Ordentlich, A., Barak, D., Stein, D., Ariel, N., and Velan, B. (1996) Aging of phosphylated human acetylcholinesterase: catalytic processes mediated by aromatic and polar residues of the active centre. *Biochem. J.* **318**, 833–840

125. Sirin, G. S., Zhou, Y., Lior-Hoffmann, L., Wang, S., and Zhang, Y. (2012) Aging mechanism of soman inhibited acetylcholinesterase. *J. Phys. Chem. B*. **116**, 12199–12207
126. Liu, N. (2015) Insecticide resistance in mosquitoes: impact, mechanisms, and research directions. *Annu. Rev. Entomol.* **60**, 537–559
127. Hemingway, J. (2000) The molecular basis of two contrasting metabolic mechanisms of insecticide resistance. *Insect Biochem. Mol. Biol.* **30**, 1009–1015
128. Marcombe, S., Poupardin, R., Darriet, F., Reynaud, S., Bonnet, J., Strode, C., Brengues, C., Yébakima, A., Ranson, H., Corbel, V., et al. (2009) Exploring the molecular basis of insecticide resistance in the dengue vector *Aedes aegypti*: a case study in Martinique Island (French West Indies). *BMC Genomics*. **10**, 494
129. Bloomquist, J. R. (1996) Ion channels as targets for insecticides. *Annu. Rev. Entomol.* **41**, 163–190
130. Soderlund, D. M., and Bloomquist, J. R. (1989) Neurotoxic actions of pyrethroid insecticides. *Annu. Rev. Entomol.* **34**, 77–96
131. Clark, J. K., Scott, J. G., Campos, F., and Bloomquist, J. R. (1995) Resistance to avermectins: extent, mechanisms, and management implications. *Annu. Rev. Entomol.* **40**, 1–30
132. Toda, S., Komazaki, S., Tomita, T., and Kono, Y. (2004) Two amino acid substitutions in acetylcholinesterase associated with pirimicarb and organophosphorous insecticide resistance in the cotton aphid, *Aphis gossypii* Glover (Homoptera: Aphididae). *Insect Mol Biol*. **13**, 549–553
133. ffrench-Constant, R. H., Pittendrigh, B., Vaughan, A., and Anthony, N. (1998) Why are there so few resistance-associated mutations in insecticide target genes? *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **353**, 1685–1693
134. Ramphul, U., Boase, T., Bass, C., Okedi, L. M., Donnelly, M. J., and Müller, P. (2009) Insecticide resistance and its association with target-site mutations in natural populations of *Anopheles gambiae* from eastern Uganda. *Trans. R. Soc. Trop. Med. Hyg.* **103**, 1121–1126
135. Liu, N., Xu, Q., Li, T., He, L., and Zhang, L. (2009) Permethrin resistance and target site insensitivity in the mosquito *Culex quinquefasciatus* in Alabama. *J. Med. Entomol.* **46**, 1424–1429
136. Weill, M., Malcolm, C., Chandre, F., Mogensen, K., Berthomieu, A., Marquine, M., and Raymond, M. (2004) The unique mutation in ace-1 giving high

insecticide resistance is easily detectable in mosquito vectors. *Insect Mol. Biol.* **13**, 1–7

137. Oakeshott, J. G., Devonshire, A. L., Claudianos, C., Sutherland, T. D., Horne, I., Campbell, P. M., Ollis, D. L., and Russell, R. J. (2005) Comparing the organophosphorus and carbamate insecticide resistance mutations in cholin- and carboxyl-esterases. *Chem. Biol. Interact.* **157**, 269–275
138. Guo, L., Xie, W., Wang, S., Wu, Q., Li, R., Yang, N., Yang, X., Pan, H., and Zhang, Y. (2014) Detoxification enzymes of *Bemisia tabaci* B and Q: biochemical characteristics and gene expression profiles. *Pest Manag. Sci.* **70**, 1588–1594
139. Hemingway, J., Coleman, M., Paton, M., McCarroll, L., Vaughan, A., and Desilva, D. (2000) Aldehyde oxidase is coamplified with the World's most common Culex mosquito insecticide resistance-associated esterases. *Insect Mol. Biol.* **9**, 93–99
140. Aïzoun, N., Aïkpon, R., Padonou, G. G., Oussou, O., Oké-Agbo, F., Gnanguenon, V., Ossè, R., and Akogbéto, M. (2013) Mixed-function oxidases and esterases associated with permethrin, deltamethrin and bendiocarb resistance in *Anopheles gambiae* sl in the south-north transect Benin, West Africa. *Parasit. Vectors*. **6**, 223
141. Zhang, Y. X., Wang, W. L., Li, M. Y., Li, S. G., and Liu, S. (2017) Identification of putative carboxylesterase and aldehyde oxidase genes from the antennae of the rice leaffolder, *Cnaphalocrocis medinalis* (Lepidoptera: Pyralidae). *J. Asia. Pac. Entomol.* **20**, 907–913
142. Low, V. L., Chen, C. D., Lee, H. L., Tan, T. K., Chen, C. F., Leong, C. S., Lim, Y. A. L., Lim, P. E., Norma-Rashid, Y., and Sofian-Azirun, M. (2013) Enzymatic characterization of insecticide resistance mechanisms in field populations of Malaysian *Culex quinquefasciatus* say (Diptera: Culicidae). *PLoS One*. **8**, e79928
143. Pavlidi, N., Vontas, J., and Van Leeuwen, T. (2018) The role of glutathione S-transferases (GSTs) in insecticide resistance in crop pests and disease vectors. *Curr. Opin. Insect Sci.* **27**, 97–102
144. Yan, S., Cui, F., and Qiao, C. (2009) Structure, function and applications of carboxylesterases from insects for insecticide resistance. *Protein Pept. Lett.* **16**, 1181–1188

145. Ochomo, E., Bayoh, M. N., Brogdon, W. G., Gimnig, J. E., Ouma, C., Vulule, J. M., and Walker, E. D. (2013) Pyrethroid resistance in *Anopheles gambiae* ss and *Anopheles arabiensis* in western Kenya: phenotypic, metabolic and target site characterizations of three populations. *Med. Vet. Entomol.* **27**, 156–164

146. Coppin, C. W., Jackson, C. J., Sutherland, T., Hart, P. J., Devonshire, A. L., Russell, R. J., and Oakeshott, J. G. (2012) Testing the evolvability of an insect carboxylesterase for the detoxification of synthetic pyrethroid insecticides. *Insect Biochem. Mol. Biol.* **42**, 343–352

147. Shi, L., Hu, H., Ma, K., Zhou, D., Yu, J., Zhong, D., Fang, F., Chang, X., Hu, S., Zou, F., et al. (2015) Development of resistance to pyrethroid in *Culex pipiens pallens* population under different insecticide selection pressures. *PLoS Negl. Trop. Dis.* **9**, e0003928

148. Casida, J. E., Gammon, D. W., Glickman, A. H., and Lawrence, L. J. (1983) Mechanisms of selective action of pyrethroid insecticides. *Annu. Rev. Pharmacol. Toxicol.* **23**, 413–438

149. Alon, M., Alon, F., Nauen, R., and Morin, S. (2008) Organophosphates' resistance in the B-biotype of *Bemisia tabaci* (Hemiptera: Aleyrodidae) is associated with a point mutation in an ace1-type acetylcholinesterase and overexpression of carboxylesterase. *Insect Biochem. Mol. Biol.* **38**, 940–949

150. Zhu, Y. C., Dowdy, A. K., and Baker, J. E. (1999) Differential mRNA expression levels and gene sequences of a putative carboxylesterase-like enzyme from two strains of the parasitoid *Anisopteromalus calandrae* (Hymenoptera: Pteromalidae). *Insect Biochem. Mol. Biol.* **29**, 417–425

151. Lee, S. H., and Clark, J. M. (1998) Permethrin carboxylesterase functions as nonspecific sequestration proteins in the hemolymph of colorado potato beetle. *Pestic. Biochem. Physiol.* **62**, 51–63

152. Devonshire, a. L. (1998) The evolution of insecticide resistance in the peach-potato aphid, *Myzus persicae*. *Philos. Trans. R. Soc. B Biol. Sci.* **353**, 1677–1684

153. Devonshire, A. L., Searle, L. M., and Moores, G. D. (1986) Quantitative and qualitative variation in the mRNA for carboxylesterases in insecticide-susceptible and resistant *Myzus persicae* (Sulz). *Insect Biochem.* **16**, 659–665

154. Li, C. X., Dong, Y. D., Song, F. L., Zhang, X. L., Gu, W. D., and Zhao, T. Y. (2009) Company amplification of estalpha2/estbeta2 and correlation between

esterase gene copy number and resistance to insecticides in the field *Culex pipiens pallens* strains collected from Beijing, China. *J. Med. Entomol.* **46**, 539–545

155. Field, L. M., and Devonshire, A. L. (1998) Evidence that the E4 and FE4 esterase genes responsible for insecticide resistance in the aphid *Myzus persicae* (Sulzer) are part of a gene family. *Biochem. J.* **330**, 169–173
156. Qiao, C. L., and Raymond, M. (1995) The same esterase B1 haplotype is amplified in insecticide-resistant mosquitoes of the *Culex pipiens* complex from the Americas and China. *Heredity (Edinb).* **74**, 339–345
157. Mouches, C., Pauplin, Y., Agarwal, M., Lemieux, L., Herzog, M., Abadon, M., Beyssat-Arnaouty, V., Hyrien, O., de Saint Vincent, B. R., and Georghiou, G. P. (1990) Characterization of amplification core and esterase B1 gene responsible for insecticide resistance in Culex. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 2574–2578
158. Mouches, C., Pasteur, N., Berge, J., Hyrien, O., Raymond, M., de Saint Vincent, B. R., de Silvestri, M., Georghiou, G. P., Mouchès, C., Pasteur, N., et al. (1986) Amplification of an esterase gene is responsible for insecticide resistance in a California Culex mosquito. *Science (80-.).* **233**, 778–780
159. Yang, X. Q., Liu, J. Y., Li, X. C., Chen, M. H., and Zhang, Y. L. (2014) Key amino acid associated with acephate detoxification by *Cydia pomonella* carboxylesterase based on molecular dynamics with alanine scanning and site-directed mutagenesis. *J. Chem. Inf. Model.* **54**, 1356–1370
160. Cui, F., Lin, Z., Wang, H., Liu, S., Chang, H., Reeck, G., Qiao, C., Raymond, M., and Kang, L. (2011) Two single mutations commonly cause qualitative change of nonspecific carboxylesterases in insects. *Insect Biochem. Mol. Biol.* **41**, 1–8
161. Newcomb, R. D., Campbell, P. M., Ollis, D. L., Cheah, E., Russell, R. J., and Oakeshott, J. G. (1997) A single amino acid substitution converts a carboxylesterase to an organophosphorus hydrolase and confers insecticide resistance on a blowfly. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 7464–7468
162. Fraser, N. J., Liu, J. W., Mabbitt, P. D., Correy, G. J., Coppin, C. W., Lethier, M., Perugini, M. A., Murphy, J. M., Oakeshott, J. G., Weik, M., et al. (2016) Evolution of protein quaternary structure in response to selective pressure for increased thermostability. *J. Mol. Biol.* **428**, 2359–2371
163. Heidari, R., Devonshire, A. L., Campbell, B. E., Bell, K. L., Dorrian, S. J., Oakeshott, J. G., and Russell, R. J. (2004) Hydrolysis of organophosphorus

insecticides by in vitro modified carboxylesterase E3 from *Lucilia cuprina*. *Insect Biochem. Mol. Biol.* **34**, 353–363

164. Devonshire, A. L., Heidari, R., Bell, K. L., Campbell, P. M., Campbell, B. E., Odgers, W. A., Oakeshott, J. G., and Russell, R. J. (2003) Kinetic efficiency of mutant carboxylesterases implicated in organophosphate insecticide resistance. *Pestic. Biochem. Physiol.* **76**, 1–13
165. Hartley, C. J., Newcomb, R. D., Russell, R. J., Yong, C. G., Stevens, J. R., Yeates, D. K., La Salle, J., and Oakeshott, J. G. (2006) Amplification of DNA from preserved specimens shows blowflies were preadapted for the rapid evolution of insecticide resistance. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 8757–8762
166. Claudianos, C., Russell, R. J., and Oakeshott, J. G. (1999) The same amino acid substitution in orthologous esterases confers organophosphate resistance on the house fly and a blowfly. *Insect Biochem. Mol. Biol.* **29**, 675–686
167. de Carvalho, R. A., Torres, T. T., and de Azeredo-Espin, A. M. L. (2006) A survey of mutations in the *Cochliomyia hominivorax* (Diptera: Calliphoridae) esterase E3 gene associated with organophosphate resistance and the molecular identification of mutant alleles. *Vet. Parasitol.* **140**, 344–351
168. McKenzie, J. A., and O'Farrell, K. (1993) Modification of developmental instability and fitness: malathion-resistance in the Australian sheep blowfly, Lucilia cuprina. *Genetica*. **89**, 67–76
169. Clarke, G. M., Yen, J. L., and McKenzie, J. A. (2000) Wings and bristles: character specificity of the asymmetry phenotype in insecticide-resistant strains of *Lucilia cuprina*. *Proc. R. Soc. B Biol. Sci.* **267**, 1815–1818
170. Davies, A. G., Game, A. Y., Chen, Z., Williams, T. J., Goodall, S., Yen, J. L., McKenzie, J. A., and Batterham, P. (1996) Scalloped wings is the *Lucilia cuprina* Notch homologue and a candidate for the modifier of fitness and asymmetry of diazinon resistance. *Genetics*. **143**, 1321–1337
171. Wang, R., Zhang, H. M., Deng, L., Xiao, Q. Q., Wang, Q. X., Liu, G. N., Han, Z. J., and Wu, M. (2017) Carboxylesterase Iscare1 mediates insecticide resistance in *Laodelphax striatellus* (Hemiptera: Delphacidae). *Acta Entomol. Sin.* **60**, 1006–1012
172. Galego, L. G. C., Ceron, C. R., and Carareto, C. M. a (2006) Characterization of esterases in a Brazilian population of *Zaprionus indianus* (Diptera:

Drosophilidae). *Genetica*. **126**, 89–99

173. Jayawardena, K. G. I., Karunaratne, S. H. P. P., Ketterman, A. J., and Hemingway, J. (1994) Determination of the role of elevated B2 esterase in insecticide resistance in *Culex quinquefasciatus* (Diptera: Culicidae) from studies on the purified enzyme. *Bull. Entomol. Res.* **84**, 39–43
174. Wang, L. L., Huang, Y., Lu, X. P., Jiang, X. Z., Smagghe, G., Feng, Z. J., Yuan, G. R., Wei, D., and Wang, J. J. (2015) Overexpression of two α-esterase genes mediates metabolic resistance to malathion in the oriental fruit fly, *Bactrocera dorsalis* (Hendel). *Insect Mol. Biol.* **24**, 467–479
175. Goh, D. K. S., Anspaugh, D. D., Motoyama, N., Rock, G. C., and Roe, R. M. (1995) Isolation and characterization of an insecticide-resistance-associated esterase in the tobacco budworm *Heliothis virescens* (F.). *Pestic. Biochem. Physiol.* **51**, 192–204
176. Rosano, G. L., and Ceccarelli, E. A. (2014) Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front. Microbiol.* **5**, 172
177. Hannig, G., and Makrides, S. C. (1998) Strategies for optimizing heterologous protein expression in *Escherichia coli*. *Trends Biotechnol.* **16**, 54–60
178. Baneyx, F. (1999) Recombinant protein expression in *Escherichia coli*. *Curr. Opin. Biotechnol.* **10**, 411–421
179. Lan, W. S., Cong, J., Jiang, H., Jiang, S. R., and Qiao, C. L. (2005) Expression and characterization of carboxylesterase E4 gene from peach-potato aphid (*Myzus persicae*) for degradation of carbaryl and malathion. *Biotechnol. Lett.* **27**, 1141–1146
180. Li, Y., Liu, J., Lu, M., Ma, Z., Cai, C., Wang, Y., and Zhang, X. (2016) Bacterial expression and kinetic analysis of carboxylesterase 001D from *Helicoverpa armigera*. *Int. J. Mol. Sci.* **17**, 493
181. Tokmakov, A. A., Kurotani, A., Takagi, T., Toyama, M., Shirouzu, M., Fukami, Y., and Yokoyama, S. (2012) Multiple post-translational modifications affect heterologous protein synthesis. *J. Biol. Chem.* **287**, 27106–27116
182. Lund, P. A. (2001) Microbial molecular chaperones. *Adv. Microb. Physiol.* **44**, 93–140
183. Hartl, F. U. (1996) Molecular chaperones in cellular protein folding. *Nature*. **381**, 571–580
184. Hatahet, F., Boyd, D., and Beckwith, J. (2014) Disulfide bond formation in

prokaryotes: history, diversity and design. *Biochim. Biophys. Acta*. **1844**, 1402–1414

185. Martínez-Alonso, M., González-Montalbán, N., García-Fruitós, E., and Villaverde, A. (2009) Learning about protein solubility from bacterial inclusion bodies. *Microb. Cell Fact.* **8**, 4
186. Prasad, S., Khadatare, P. B., and Roy, I. (2011) Effect of chemical chaperones in improving the solubility of recombinant proteins in *Escherichia coli*. *Appl. Environ. Microbiol.* **77**, 4603–4609
187. de Marco, A. (2012) Recent contributions in the field of the recombinant expression of disulfide bonded protein in bacteria. *Microb. Cell Fact.* **11**, 129
188. de Marco, A. (2009) Strategies for successful recombinant expression of disulfide bond-dependent proteins in *Escherichia coli*. *Microb. Cell Fact.* **8**, 26
189. Chen, J., Song, J. L., Zhang, S., Wang, Y., Cui, D. F., and Wang, C. C. (1999) Chaperone activity of DsbC. *J. Biol. Chem.* **274**, 19601–19605
190. Terpe, K. (2006) Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Appl. Microbiol. Biotechnol.* **72**, 211–222
191. Nielsen, H. (2017) Predicting secretory proteins with SignalP. *Methods Protoc.* **1611**, 59–73
192. Ferre, F., and Clote, P. (2006) DiANNA 1.1: an extension of the DiANNA web server for ternary cysteine classification. *Nucleic Acids Res.* **34**, W182–W185
193. Smialowski, P., Doose, G., Torkler, P., Kaufmann, S., and Frishman, D. (2012) PROSO II - a new method for protein solubility prediction. *FEBS J.* **279**, 2192–2200
194. Slabinski, L., Jaroszewski, L., Rychlewski, L., Wilson, I. A., Lesley, S. A., and Godzik, A. (2007) XtalPred: a web server for prediction of protein crystallizability. *Bioinformatics*. **23**, 3403–3405
195. Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., and Sayers, E. W. (2016) GenBank. *Nucleic Acids Res.* **44**, D67–D72
196. The UniProt Consortium (2017) UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* **45**, D158–D169
197. Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., and Smith, H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods*. **6**, 343–345

198. Leland, J., and Gore, J. (2017) Microbial control of insect and mite pests of cotton. in *Microbial control of insect and mite pests: from theory to practice* (Lacey, Lawrence, A. ed), pp. 185–197, Academic Press, Amsterdam
199. Magaña, C., Hernández-Crespo, P., Brun-Barale, A., Couso-Ferrer, F., Bride, J.-M., Castañera, P., Feyereisen, R., and Ortego, F. (2008) Mechanisms of resistance to malathion in the medfly *Ceratitis capitata*. *Insect Biochem. Mol. Biol.* **38**, 756–762
200. Vaughan, A., Hawkes, N., and Hemingway, J. (1997) Co-amplification explains linkage disequilibrium of two mosquito esterase genes in insecticide-resistant *Culex quinquefasciatus*. *Biochem. J.* **325**, 359–365
201. Nene, V., Wortman, J. R., Lawson, D., Haas, B., Kodira, C., Tu, Z., Loftus, B., Xi, Z., Megy, K., Grabherr, M., et al. (2007) Genome sequence of *Aedes aegypti*, a major arbovirus vector. *Science (80-.).* **316**, 1718–1723
202. Holt, R. A., Subramanian, G. M., Halpern, A., Sutton, G. G., Charlab, R., Nusskern, D. R., Wincker, P., Clark, A. G., Ribeiro, J. M. C., Wides, R., et al. (2002) The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science*. **298**, 129–149
203. Pan, Y., Guo, H., and Gao, X. (2009) Carboxylesterase activity, cDNA sequence, and gene expression in malathion susceptible and resistant strains of the cotton aphid, *Aphis gossypii*. *Comp. Biochem. Physiol. B. Biochem. Mol. Biol.* **152**, 266–270
204. Zhang, J., Ge, P., Li, D., Guo, Y., Zhu, K. Y., Ma, E., and Zhang, J. (2015) Two homologous carboxylesterase genes from *Locusta migratoria* with different tissue expression patterns and roles in insecticide detoxification. *J. Insect Physiol.* **77**, 1–8
205. Small, G. J., and Hemingway, J. (2000) Molecular characterization of the amplified carboxylesterase gene associated with organophosphorus insecticide resistance in the brown planthopper, *Nilaparvata lugens*. *Insect Mol. Biol.* **9**, 647–653
206. Field, L. M., Williamson, M. S., Moores, G. D., and Devonshire, A. L. (1993) Cloning and analysis of the esterase genes conferring insecticide resistance in the peach-potato aphid, *Myzus persicae* (Sulzer). *Biochem. J.* **294**, 569–574
207. Hernandez, R., He, H., Chen, A. C., Waghela, S. D., Wayne Ivie, G., George, J. E., and Gale Wagner, G. (2000) Identification of a point mutation in an esterase

gene in different populations of the southern cattle tick, *Boophilus microplus*. *Insect Biochem. Mol. Biol.* **30**, 969–977

208. Cui, F., Qu, H., Cong, J., Liu, X. L., and Qiao, C. L. (2007) Do mosquitoes acquire organophosphate resistance by functional changes in carboxylesterases? *FASEB J.* **21**, 3584–3591
209. Jaroszewski, L., Slabinski, L., Wooley, J., Deacon, A. M., Lesley, S. A., Wilson, I. A., and Godzik, A. (2008) Genome pool strategy for structural coverage of protein families. *Structure*. **16**, 1659–1667
210. Jahandideh, S., Jaroszewski, L., and Godzik, A. (2014) Improving the chances of successful protein structure determination with a random forest classifier. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **70**, 627–635
211. Grabski, A. C. (2009) Advances in preparation of biological extracts for protein purification. in *Methods in Enzymology*, 2nd Ed. (Burgess, R. R., and Deutscher, M. P. eds), pp. 285–303, Academic Press, **463**, 285–303
212. Demain, A. L., and Vaishnav, P. (2009) Production of recombinant proteins by microbes and higher organisms. *Biotechnol. Adv.* **27**, 297–306
213. Wang, R. F., and Kushner, S. R. (1991) Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in *Escherichia coli*. *Gene*. **100**, 195–199
214. Burgess, R. R. (2009) Refolding solubilized inclusion body proteins. in *Methods in Enzymology*, pp. 259–282, Academic Press, **463**, 259–282
215. Yang, Z., Zhang, L., Zhang, Y., Zhang, T., Feng, Y., Lu, X., Lan, W., Wang, J., Wu, H., Cao, C., et al. (2011) Highly efficient production of soluble proteins from insoluble inclusion bodies by a two-step-denaturing and refolding method. *PLoS One*. **6**, e22981
216. Zhu, Y. C., Dowdy, A. K., and Baker, J. E. (1999) Detection of single-base substitution in an esterase gene and its linkage to malathion resistance in the parasitoid *Anisopteromalus calandrae* (Hymenoptera: Pteromalidae). *Pestic. Sci.* **55**, 398–404
217. Liu, Y., Zhang, H., Qiao, C., Lu, X., and Cui, F. (2011) Correlation between carboxylesterase alleles and insecticide resistance in *Culex pipiens* complex from China. *Parasit. Vectors*. **4**, 236
218. Zhang, H. H., Meng, F., Qiao, C. C., Cui, F., Cui, F., Raymond, M., Qiao, C. C., Cheikh, R. Ben, Berticat, C., Berthomieu, A., et al. (2012) Identification of

resistant carboxylesterase alleles in *Culex pipiens* complex via PCR-RFLP. *Parasit. Vectors*. **5**, 209

219. Hemingway, J., and Karunaratne, S. H. (1998) Mosquito carboxylesterases: a review of the molecular biology and biochemistry of a major insecticide resistance mechanism. *Med. Vet. Entomol.* **12**, 1–12
220. Uberto, R., and Moomaw, E. W. (2013) Protein similarity networks reveal relationships among sequence, structure, and function within the Cupin superfamily. *PLoS One*. **8**, e74477
221. Atkinson, H. J., Morris, J. H., Ferrin, T. E., and Babbitt, P. C. (2009) Using sequence similarity networks for visualization of relationships across diverse protein superfamilies. *PLoS One*. **4**, e4345
222. Gerlt, J. A., Bouvier, J. T., Davidson, D. B., Imker, H. J., Sadkhin, B., Slater, D. R., and Whalen, K. L. (2015) Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A web tool for generating protein sequence similarity networks. *Biochim. Biophys. Acta*. **1854**, 1019–1037
223. Brown, S. D., and Babbitt, P. C. (2012) Inference of functional properties from large-scale analysis of enzyme superfamilies. *J. Biol. Chem.* **287**, 35–42
224. Han, Q., Wong, D. M., Robinson, H., Ding, H., Lam, P. C. H., Totrov, M. M., Carlier, P. R., and Li, J. (2017) Crystal structure of acetylcholinesterase catalytic subunits of the malaria vector *Anopheles gambiae*. *Insect Sci.* 10.1111/1744-7917.12450
225. Ycas, M. (1974) On earlier states of the biochemical system. *J. Theor. Biol.* **44**, 145–160
226. . Newton, M. S., Arcus, V. L., Gerth, M. L., and Patrick, W. M. (2018) Enzyme evolution: innovation is easy, optimization is complicated. *Curr. Opin. Struct. Biol.* **48**, 110–116
227. Carroll, S. B. (2005) Evolution at two levels: on genes and form. *PLoS Biol.* **3**, e245
228. Hoekstra, H. E., and Coyne, J. A. (2007) The locus of evolution: evo devo and the genetics of adaptation. *Evolution (N. Y).* **61**, 995–1016
229. Tangwancharoen, S., Moy, G. W., and Burton, R. S. (2018) Multiple modes of adaptation: regulatory and structural evolution in a small heat shock protein gene. *Mol. Biol. Evol.* **35**, 2110–2119
230. Juneja, P., Quinn, A., and Jiggins, F. M. (2016) Latitudinal clines in gene

expression and cis-regulatory element variation in *Drosophila melanogaster*. *BMC Genomics*. **17**, 981

231. Campbell, P. M., Harcourt, R. L., Crone, E. J., Claudianos, C., Hammock, B. D., Russell, R. J., and Oakeshott, J. G. (2001) Identification of a juvenile hormone esterase gene by matching its peptide mass fingerprint with a sequence from the Drosophila genome project. *Insect Biochem. Mol. Biol.* **31**, 513–520
232. Kamita, S. G., and Hammock, B. D. (2010) Juvenile hormone esterase: biochemistry and structure. *J. Pestic. Sci.* **35**, 265–274
233. Kontogiannatos, D., Michail, X., and Kourti, A. (2011) Molecular characterization of an ecdysteroid inducible carboxylesterase with GQSCG motif in the corn borer, *Sesamia nonagrioides*. *J. Insect Physiol.* **57**, 1000–1009
234. Zhu, L., Yin, T. Y., Sun, D., Liu, W., Zhu, F., Lei, C. L., and Wang, X. P. (2017) Juvenile hormone regulates the differential expression of putative juvenile hormone esterases via methoprene-tolerant in non-diapause-destined and diapause-destined adult female beetle. *Gene*. **627**, 373–378
235. El-Sheikh, E. A. (2015) Characterization and kinetics of juvenile hormone esterase from *Spodoptera littoralis* (Boisd.) and *S. frugiperda* (J. E. Smith) (Lepidoptera: Noctuidae). *Egypt. J. Biol. Pest Control*. **25**, 625–630
236. Elayidam, U. G., and Muraleedharen, D. (2008) Identification and partial characterization of juvenile hormone esterase from cotton pest *Dysdercus cingulatus*. *Indian J. Biochem. Biophys.* **45**, 121–125
237. Kamita, S. G., Samra, A. I., Liu, J.-Y., Cornel, A. J., and Hammock, B. D. (2011) Juvenile hormone (JH) esterase of the mosquito *Culex quinquefasciatus* is not a target of the JH analog insecticide methoprene. *PLoS One*. **6**, e28392
238. Kontogiannatos, D., Swevers, L., Maenaka, K., Park, E. Y., Iatrou, K., and Kourti, A. (2013) Functional characterization of a juvenile hormone esterase related gene in the moth *Sesamia nonagrioides* through RNA interference. *PLoS One*. **8**, e73834
239. Valaitis, A. P. (1991) Characterization of hemolymph juvenile hormone esterase from *Lymantria dispar*. *Insect Biochem.* **21**, 583–595
240. Feng, Q.-L., Ladd, T. R., Tomkins, B. L., Sundaram, M., Sohi, S. S., Retnakaran, A., Davey, K. G., and Palli, S. R. (1999) Spruce budworm (*Choristoneura fumiferana*) juvenile hormone esterase: hormonal regulation, developmental expression and cDNA cloning. *Mol. Cell. Endocrinol.* **148**, 95–108

241. Wogulis, M., Wheelock, C. E., Kamita, S. G., Hinton, A. C., Whetstone, P. A., Hammock, B. D., and Wilson, D. K. (2006) Structural studies of a potent insect maturation inhibitor bound to the juvenile hormone esterase of *Manduca sexta*. *Biochemistry*. **45**, 4045–4057

242. Campbell, P. M., Oakeshott, J. G., and Healy, M. J. (1998) Purification and kinetic characterisation of juvenile hormone esterase from *Drosophila melanogaster*. *Insect Biochem. Mol. Biol.* **28**, 501–515

243. Kamita, S. G., Wogulis, M. D., Law, C. S., Morisseau, C., Tanaka, H., Huang, H., Wilson, D. K., and Hammock, B. D. (2010) Function of phenylalanine 259 and threonine 314 within the substrate binding pocket of the juvenile hormone esterase of *Manduca sexta*. *Biochemistry*. **49**, 3733–3742

244. Kamita, S. G., Hinton, A. C., Wheelock, C. E., Wogulis, M. D., Wilson, D. K., Wolf, N. M., Stok, J. E., Hock, B., and Hammock, B. D. (2003) Juvenile hormone (JH) esterase: why are you so JH specific? *Insect Biochem. Mol. Biol.* **33**, 1261–1273

245. Leal, W. S. (2013) Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. *Annu. Rev. Entomol.* **58**, 373–391

246. Vogt, R. G., and Riddiford, L. M. (1981) Pheromone binding and inactivation by moth antennae. *Nature*. **293**, 161–163

247. Ishida, Y., and Leal, W. S. (2005) Rapid inactivation of a moth pheromone. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 14075–14079

248. Merlin, C., Rosell, G., Carot-Sans, G., François, M.-C., Bozzolan, F., Pelletier, J., Jacquin-Joly, E., Guerrero, A., and Maïbèche-Coisne, M. (2007) Antennal esterase cDNAs from two pest moths, *Spodoptera littoralis* and *Sesamia nonagrioides*, potentially involved in odourant degradation. *Insect Mol. Biol.* **16**, 73–81

249. He, P., Zhang, J., Li, Z. Q., Zhang, Y. N., Yang, K., Dong, S. L., and He, P. (2014) Functional characterization of an antennal esterase from the noctuid moth, *Spodoptera exigua*. *Arch. Insect Biochem. Physiol.* **86**, 85–99

250. Steiner, C., Bozzolan, F., Montagné, N., Maïbèche, M., and Chertemps, T. (2017) Neofunctionalization of “juvenile hormone esterase duplication” in Drosophila as an odorant-degrading enzyme towards food odorants. *Sci. Rep.* **7**, 12629

251. Younus, F., Chertemps, T., Pearce, S. L., Pandey, G., Bozzolan, F., Coppin, C.

W., Russell, R. J., Maïbèche-Coisne, M., and Oakeshott, J. G. (2014) Identification of candidate odorant degrading gene/enzyme systems in the antennal transcriptome of *Drosophila melanogaster*. *Insect Biochem. Mol. Biol.* **53**, 30–43

252. Copley, S. D. (2017) Shining a light on enzyme promiscuity. *Curr. Opin. Struct. Biol.* **47**, 167–175
253. Ferla, M. P., Brewster, J. L., Hall, K. R., Evans, G. B., and Patrick, W. M. (2017) Primordial-like enzymes from bacteria with reduced genomes. *Mol. Microbiol.* **105**, 508–524
254. Kidwell, M. G., and Lisch, D. (1997) Transposable elements as sources of variation in animals and plants. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 7704–7711
255. Wittkopp, P. J., and Kalay, G. (2012) Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nat. Rev. Genet.* **13**, 59–69
256. Levin, H. L., and Moran, J. V. (2011) Dynamic interactions between transposable elements and their hosts. *Nat. Rev. Genet.* **12**, 615–627
257. Birner-Gruenberger, R., Bickmeyer, I., Lange, J., Hehlert, P., Hermetter, A., Kollroser, M., Rechberger, G. N., and Kühnlein, R. P. (2012) Functional fat body proteomics and gene targeting reveal in vivo functions of *Drosophila melanogaster* α-Esterase-7. *Insect Biochem. Mol. Biol.* **42**, 220–229
258. Wen, D., Rivera-Perez, C., Abdou, M., Jia, Q., He, Q., Liu, X., Zyaan, O., Xu, J., Bendena, W. G., Tobe, S. S., et al. (2015) Methyl farnesoate plays a dual role in regulating Drosophila metamorphosis. *PLOS Genet.* **11**, e1005038
259. Jones, G., Jones, D., Li, X., Tang, L., Ye, L., Teal, P., Riddiford, L., Sandifer, C., Borovsky, D., and Martin, J.-R. (2010) Activities of natural methyl farnesoids on pupariation and metamorphosis of *Drosophila melanogaster*. *J. Insect Physiol.* **56**, 1456–1464
260. Tsubota, T., Shimomura, M., Ogura, T., Seino, A., Nakakura, T., Mita, K., Shinoda, T., and Shiotsuki, T. (2010) Molecular characterization and functional analysis of novel carboxyl/cholinesterases with GQSAG motif in the silkworm *Bombyx mori*. *Insect Biochem. Mol. Biol.* **40**, 100–112
261. Babtie, A., Tokuriki, N., and Hollfelder, F. (2010) What makes an enzyme promiscuous? *Curr. Opin. Chem. Biol.* **14**, 200–207
262. Kondrashov, F. A. (2005) In search of the limits of evolution. *Nat. Genet.* **37**,9-10

263. Correy, G., Zaidman, D., Carvalho, S., Mabbitt, P. D., James, P. J., Kotze, A. C., London, N., and Jackson, C. J. (2017) Overcoming insecticide resistance through computational inhibitor design. *bioRxiv*. 10.1101/161430